LA DUALITE DE LA PROBABILITE DANS L'ENSEIGNEMENT DE LA STATISTIQUE. UNE EXPERIENCE EN CLASSE DE BTS

Carranza Pablo

UNIVERSITE PARIS 7 – DENIS DIDEROT
UFR DE MATHEMATIQUES
Ecole Doctorale « Savoirs scientifiques : épistémologie, histoire des sciences et didactique des disciplines »

THESE
pour l'obtention du diplôme de
Docteur de l'UNIVERSITE PARIS 7
Spécialité : DIDACTIQUE DES MATHEMATIQUES

Présentée et soutenue publiquement le 11 Juin 2009 par
Pablo CARRANZA

LA DUALITE DE LA PROBABILITE DANS L'ENSEIGNEMENT DE LA STATISTIQUE. UNE EXPERIENCE EN CLASSE DE BTS

Directeur de thèse :
M. Alain KUZNIAK

Membres du JURY

Mme. María del Pilar Orús BAGUENA	Professeur, Université Jaume I	Rapporteur
M. Christophe HACHE	Professeur, Université Paris 7	Rapporteur
M. Alain KUZNIAK	Professeur, Université Paris 7	Directeur de Thèse
M. Jean-Claude REGNIER	Professeur, Université Lyon 2	Examinateur
M. Bernard PARZYSZ	Professeur, Université Paris 7	Président du Jury

Table de matières

Introduction 2
Chapitre I : Enquête épistémologique 11
1.1 Introduction 11
1.2 Dualité d'interprétation de la probabilité et recherche d'éléments caractéristiques 14
1.3 Blaise Pascal. Un début de la dualité 18
1.4 La Logique de Port-Royal. Les débuts de la quantification 25
1.5 Gottfried Leibniz. Vers une probabilité logique 32
1.6 Jacques Bernoulli. Le théorème limite 35
1.7 Thomas Bayes. L'inversion de la probabilité 39
1.8 Pierre-Simon Laplace. Une définition opératoire 49
1.9 John Keynes. Une approche logique 56
1.10 Bruno De Finetti. Une probabilité subjective 66
1.11 Richard von Mises. La probabilité fréquentiste 72
1.12 Karl Popper. La probabilité propensioniste 78
1.13 Conclusions 82
Chapitre II : Analyse de manuels 92
2.1 Introduction 92
2.2 L'approche des rédacteurs des programmes 94
2.3 Problématique 96
2.4 Méthodologie 99
2.5 Manuel Belin S 125
2.6 Manuel Nathan S 146
2.7 Manuel Bréal ES 162
2.8 Manuel Didier ES 174
2.9 Conclusions 186
Chapitre III : Expérimentations en BTS 193
4.1 Introduction 193
4.1 Caractéristiques générales 195
4.1 Caractéristiques spécifiques 204
4.1 Analyse de la première expérimentation 205
4.1 Analyse de deuxième expérimentation 244
4.1 Analyse de troisième expérimentation 289
Chapitre IV : Conclusions et perspectives 333
4.1 Introduction 333
4.2 Enquête épistémologique 333
4.3 Analyse de manuels 341
4.4 Expérimentations en BTS 346
Bibliographie 368
Annexes 371
Programme et Documents d'accompagnement 372
Transcription séance Jeu de pièces de monnaie 376
Transcription séance Les circuits 396
Transcription séance La bouteille 439
Feuilles de recherche. Séance les circuits 473
Questionnaire. Séance la bouteille 479

Introduction

Dans ce travail de thèse, nous nous intéressons aux possibilités de sensibiliser à la dualité de la probabilité dans l'enseignement en France. Pour cela nous avons organisé cette présentation en quatre parties. La première est consacrée à une enquête épistémologique, la deuxième à une analyse de manuels, la troisième à un ensemble d'expérimentations effectuées en BTS et finalement dans la quatrième nous donnerons quelques conclusions et perspectives qui découlent de notre travail.

Nous parlons bien de sensibilisation car notre thèse peut se décrire comme une sorte d'exploration des possibilités (et conditions) de l'enseignement de la probabilité. Notre originalité étant, que dans cette recherche la probabilité est considérée comme admettant deux interprétations, l'une fréquentiste, l'autre bayésienne. En effet, tout au long de notre travail, nous nous intéresserons aux possibilités de traitement de ces deux interprétations de la probabilité dans une classe de mathématique.

Considérer la probabilité comme étant une notion duale pourrait apparaître comme un choix personnel, un parmi d'autres. Néanmoins, nous montrerons, et cela sera l'objet du premier chapitre, que la probabilité à gardé depuis ses origines et jusqu'à nos jours deux versants interprétatifs, l'un caractérisé comme la fréquence d'apparition d'un phénomène donné, l'autre comme un degré de certitude porté sur une proposition donnée, ces deux interprétations sont à l'origine de deux méthodes ou écoles inférentielles largement reconnues dans les sociétés savantes.

Le premier chapitre est donc le résumé d'une enquête où nous nous sommes intéressés aux caractéristiques épistémologiques de cette notion, ce chapitre devient la référence épistémologique de notre synthèse. Nous y reproduisons quelques fragments d'écrits des principaux acteurs de la construction de cette notion. Dans cette présentation, nous retenons des fragments où les auteurs nous fournissent des indices de la présence des deux approches de la probabilité. De cette manière, nous avons transcrit des extraits de textes de Pascal, Leibniz, Bernoulli, Bayes, de Finetti, Popper, etc.

De l'enquête épistémologique nous avons aussi retenu une autre conclusion que celle de la dualité de la probabilité et cette autre conclusion sera déterminante pour notre thèse. Elle concerne le caractère incontournable de la dualité d'interprétation de la probabilité. En d'autres termes, nous postulons que la dualité de la probabilité sur le plan épistémologique se trouve indéfectiblement présente sur le plan didactique. De cette manière, la dualité de signifié qui est une caractéristique épistémologique, le sera aussi lors de la transposition didactique. Tout projet d'enseignement de cette notion doit, selon notre approche, intégrer les deux interprétations, la fréquentiste et la bayésienne.

Dans le premier chapitre de cette synthèse, nous chercherons donc à présenter des traces de cette dualité puis , dans le deuxième, à valider son caractère incontournable en montrant que malgré toutes les tentatives[1] de découpage de cette dualité, elle se présente inévitablement dans l'enseignement. Pour vérifier cela, nous avons effectué une étude des exercices de manuels de lycée.

Mais avant de réaliser cette étude sur les manuels, une autre tâche s'est avérée indispensable. Elle concerne la caractérisation des deux interprétations de la probabilité. En d'autres termes, si la probabilité admet deux versants, nous devons d'abord pouvoir les identifier en les différenciant le plus clairement possible l'un de l'autre. Cette caractérisation se réalise sur le plan épistémologique, avant toute transposition didactique. Ceci constitue l'objet de la seconde partie du premier chapitre, nous cherchons donc non seulement à repérer des traces de cette dualité tout au long de l'histoire de la probabilité mais aussi à trouver des éléments permettant d'identifier les deux interprétations.

Il nous semble nécessaire de préciser notre position sur cette intention de caractériser des interprétations de la probabilité. Nous admettons qu'à un type de contexte donné corresponde une interprétation de la probabilité, et vice-versa. En d'autres terme, il y a, selon notre approche, une sorte de bijection entre deux ensembles : l'ensemble de contextes et l'ensemble des interprétations de la probabilité. La caractérisation proposée dans cette synthèse a pour fonction de relier ces deux ensembles. De cette manière, en pointant certaines caractéristiques du problème, il est possible de reconnaître l'interprétation de la probabilité associée. Une conséquence de cette approche est que l'interprétation associée à un problème donné devient une tâche (relativement) objective. En effet, en identifiant les éléments caractéristiques du problème nous pouvons lui associer une interprétation de la probabilité. De cette manière, l'interprétation n'est pas un « attribut » dépendant de l'observateur du problème mais une conséquence des caractéristiques de la situation. Si les choses se passaient autrement, en d'autres termes, si il était possible d'interpréter librement un calcul de la probabilité, il serait impossible de parler d'erreurs d'interprétation, tout simplement parce que chaque observateur

[1] Le programme français en vigueur l'année de nos expérimentations établit que le terme probabilité doit être associé à la stabilisation de fréquences. Aucune mention n'est faite à la notion bayésienne.

pourrait associer librement à un contexte donné l'interprétation de la probabilité la plus convenable à ses yeux, approche de laquelle nous nous distancions clairement.

Cette sorte de bijection entre contexte et interprétation nous sera d'une grande utilité sur le plan didactique. En effet, lorsque nous nous intéresserons à l'enseignement d'une interprétation particulière, il sera nécessaire de considérer le type de contextes auquel l'interprétation visée s'associe puis de concevoir des situations répondant à ces caractéristiques. En termes issus de la théorie des champs conceptuels de G. Vergnaud (Vergnaud, 1990), il s'agit d'identifier le genre de situations auxquelles chaque interprétation de la probabilité donne réponse. Dans le troisième chapitre, lorsque nous mènerons des expérimentations visant à la sensibilisation aux deux interprétations, nous utiliserons les éléments caractéristiques proposés lors du premier chapitre pour concevoir les problèmes de nos situations. De cette manière, nous envisageons la bijection dans le sens interprétation visée vers le contexte : pour sensibiliser à une interprétation donnée, nous cherchons son contexte approprié. Dans le deuxième chapitre nous regarderons la bijection réciproque, c'est-à-dire du contexte vers l'interprétation. En effet, c'est à partir de l'observation de quelques caractéristiques dans des exercices que nous déterminerons l'interprétation qui leur est associé. De cette manière, l'analyse des manuels pourrait se résumer à une recherche sur la présence de contextes « bayésiens » et « fréquentistes ». En effet, en confirmant la présence de contextes des deux types nous validerons le caractère incontournable de la probabilité lors de sa transposition didactique dans une des étapes de cette transposition, celle des exercices des manuels. Le lecteur verra ainsi que la caractérisation des contextes associés à chacune des interprétations devient fondamentale dans notre projet, elle nous sert dans le deuxième chapitre pour valider notre hypothèse et dans le troisième pour nous outiller lors de nos expérimentations.

La Figure 1 est une représentation de la bijection entre contextes et interprétations de la probabilité, l'association d'un problème à son interprétation est donnée par les éléments caractéristiques. Dans le chapitre II, lorsque nous repérerons l'interprétation associée à un problème, nous parcourons le graphique de gauche à droite, en effet, c'est en identifiant quelques caractéristiques dans les exercices que nous leur associerons une interprétation de la probabilité sous-jacente. Et c'est précisément en trouvant des exercices associés les uns à l'interprétation fréquentiste, les autres à la bayésienne que nous confirmons le caractère incontournable de la dualité de la probabilité. Dans le chapitre III, nous parcourrons la figure dans le sens opposé, de droite à gauche, en choisissant de sensibiliser à une interprétation donnée, nous dessinons une situation-problème répondant à ses éléments caractéristiques associés.

Application d'éléments en l'association entre problèmes et interprétations

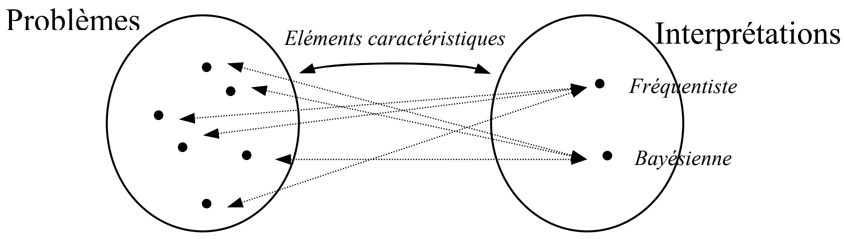


Figure 1

Ces éléments caractéristiques, au nombre de quatre, ont été dégagés lors de notre enquête épistémologique. Ils sont tous reliés, et pris ensemble, tendent à fournir un ensemble cohérent pour l'observation et un instrument pour l'identification de l'interprétation de la probabilité sous-jacente à un problème donné. Ces éléments caractéristiques, présentés pour la première fois dans cette thèse, nous semblent nécessaires à la tâche d'identification d'une interprétation, néanmoins, et très probablement ils devront être réexaminer avec soin, à fin de les compléter et mieux les cerner, cela sera l'objet des perspectives ouvertes par cette thèse.

Rappelons que la conclusion de l'enquête épistémologique affirme d'une part la dualité de l'interprétation de la probabilité, et d'autre part son caractère indissociable, même lors de sa transposition didactique. La deuxième partie de cette conclusion devient notre hypothèse didactique. Si les interprétations de la probabilité sont indissociables, leur enseignement commun devient incontournable. C'est ainsi que nous nous sommes intéressés à la sensibilisation aux deux interprétations de la probabilité.

La validation du caractère incontournable de cette dualité sera abordé par l'étude des manuels, pour cela nous avons pris quatre manuels de lycée français, deux de la filière Scientifique et deux autres de la filière Economie et Sciences Sociales, tous les quatre correspondant à la classe de Première. Le choix de cette année est la conséquence des directives officielles, c'est en la classe de Première que les rédacteurs des programmes ont choisi d'introduire le concept de probabilité, tant dans sa dimension interprétative que calculatoire. Le concept est bien évidemment repris dans les années suivantes, mais c'est en la classe de Première qu'il est défini pour les années qui suivent. Nous nous sommes donc centrés sur les manuels de cette année pour valider la présence de contextes tant bayésiens que fréquentistes dans les exercices proposés aux élèves.

Les conclusions à tirer de notre étude sur les manuels restent limitées, malheureusement. D'une part, parce que nous avons restreint le nombre de manuels analysés à quatre et cela empêche tout type de validation lors d'une éventuelle tentative de généralisation. Et d'autre part, parce que cette étude ne permet de repérer que la présence de contextes associés aux deux types d'interprétations et dans aucun cas elle ne nous permet d'inférer le traitement effectivement donné en classe. En effet, cette étude, basée exclusivement sur l'espace de travail potentiel () proposé par les exercices de quatre manuels, ne nous renseigne pas sur les possibles transformations qu'enseignants et élèves peuvent effectuer lors du travail en classe. Cette étude nous renseigne uniquement sur des interprétations de la probabilité sous-jacentes aux exercices auxquels seraient confrontés les élèves lors de leur passage au lycée.

Malgré ces limitations, l'information tirée de l'étude nous semble non négligeable. En effet, si nous admettons que ces exercices ne subissent pas de modifications substantielles de contexte lors de leur traitement en classe et que, de plus, les manuels d'une même filière tendent à se rapprocher, une éventuelle validation de la présence des contextes bayésiens et fréquentistes dans ces quatre manuels nous suggèrera que la dualité de signifiés de la probabilité existe et ceci malgré les directives officielles. De cette manière, les deux interprétations se manifesteraient en classe, l'une, la fréquentiste, arriverait à avoir un statut officiel (institutionnalisation) tandis que l'autre, la bayésienne, resterait « cachée » et alors tout type de conceptualisation associée à cette interprétation resterait entièrement à la charge des élèves, sans possibilité de mise en commun.

Telle que nous l'avons signalée ci-dessus, l'étude des manuels vise à valider la présence des deux interprétations dans les exercices des manuels. Cette validation nous semble intéressante en soi, néanmoins il nous est apparu important aussi d'avancer un peu plus sur les exercices en tentant d'identifier des profils d'exercices, afin de nous approcher d'une caractérisation des exercices du chapitre probabilité de la classe de Première. Avec cette tentative de caractérisation, nous poursuivons deux grands objectifs. Premièrement, comprendre de quelle manière les manuels parviennent à résoudre ce qui semblerait une contradiction donnée par une directive non suivie. Cette contradiction pouvant se résumer à l'obligation donnée, d'une part, par les directives officielles de devoir associer la probabilité exclusivement à la notion fréquentiste et , d'autre part, dans les faits de proposer aux élèves des exercices tant bayésiens que fréquentistes. Deuxièmement, et cette fois-ci, cela concerne nos expérimentations, la caractérisation des exercices nous permettra de disposer d'information concernant les élèves. En effet, en identifiant des profils d'exercices nous pourrions inférer le genre de situations auxquelles seraient habitués les élèves du BTS avec qui nous effectuerons nos expérimentations. Cette information nous est apparue très riche au moment de concevoir nos situations-problèmes, elle nous aidera, entre autres, à comprendre les habitudes non seulement des élèves mais aussi celles de leur enseignant.

Pour caractériser les exercices en termes de profils nous avons introduit un ensemble de variables dont le lecteur trouvera leur détail dans la première partie du deuxième chapitre. Avec ces variables nous cherchons à profiler les exercices des manuels observés en visant d'une part à comprendre une éventuelle contradiction entre l'interprétation soutenue par le programme officiel et les interprétations sous-jacentes aux exercices, et d'autre part à nous outiller pour les expérimentations.

Pour cette recherche de profils d'exercices nous avons retenu la méthode d'analyse implicative, méthode qui s'est avérée appropriée à notre problématique. En effet, nous représenterons la caractérisation envisagée sous la forme d'implications. Les plus élémentaires considéreront les types de contextes dans la premisse d'une implication, par exemple nous nous intéressions à des questions du genre « si un exercice est de nature bayésienne, alors…. » ou « si un exercice est de nature fréquentiste, alors… ». Dans ce sens, les implications sont plus proches de notre but que par exemple l'étude des corrélations. Bien entendu, étant donnée la variabilité entre exercices, il est impossible de valider des implications logiques, le seul moyen possible sera une validation statistique sous la forme de quasi-implications.

Les conclusions seront données au lecteur pour chaque manuel et de manière plus générale sur la totalité des manuels à la fin du Chapitre II. Nous pouvons déjà signaler ici que l'étude nous a confirmé le caractère dual de la probabilité lors de sa transposition didactique. Avant de faire le traitement en classe et de connaître le sens finalement retenu par les élèves, cette étude nous a indiqué que la dualité d'interprétation de la probabilité est bien présente dans les quatre manuels observés. En effet, nous avons pu constater la présence de contextes les uns de nature bayésienne les autres de nature fréquentiste. D'ailleurs, l'autre objectif poursuivi, celui de profiler les exercices, nous a aidé d'une part à comprendre de quelle manière les manuels parviennent à résoudre le conflit entre la dualité incontournable et les directives officielles, et d'autre part à identifier quelques tendances dans les exercices. Ces tendances nous ont permis d'inférer le genre d'exercices auxquels auraient été confrontés les élèves du BTS, lors de leur passage par le lycée.

Toute cette information tirée de l'analyse de manuels est reprise dans l'étape suivante, celle de nos expérimentations. C'est dans le troisième chapitre, lorsque nous nous intéresserons à quelques expérimentations en classe que nous reprendrons les conclusions de l'analyse des manuels. Dans ce troisième chapitre, nous reprendrons aussi les conclusions de l'enquête épistémologique, en particulier les éléments caractéristiques de chacune des interprétations. En général, le lecteur verra que les éléments caractéristiques du premier chapitre et les informations tirées des manuels se sont traduits en variables didactiques et/ou en conditions à remplir par nos situations-problémes.

Nous disposions donc, grâce aux deux premiers chapitres, d'informations fondamentales pour concevoir nos situations. Parmi d'autres, une concerne les registres sémiotiques. En effet, dans les deux premiers chapitres, nous avons progressivement retenu deux dimensions pour la probabilité, l'une calculatoire, l'autre sémantique. La première porte sur la valeur numérique d'une probabilité, la deuxième sur le signifié attribué à cette valeur, chacune se développant dans des registres sémiotiques différentes. La dimension calculatoire le fait dans le registre numérique, la dimension sémantique dans le registre langagier. Cette distinction devient importante si l'on s'intéresse aux interprétations de la probabilité, car elle nous renseigne sur le genre de registre que nous devons pointer en fonction de notre l'intérêt. Pour dégager les interprétations de la probabilité, nous devons nous intéresser (et favoriser chez les élèves) à la possibilité d'expression dans le registre langagier, seul registre à disposer d'outils pour qu'une interprétation de la probabilité soit précisée.

En effet, ni le registre numérique, ni le registre symbolique ne permettent à un interlocuteur de préciser le sens attribué à une probabilité. Par exemple, l'expression $P(A) = 1/3$ réunit deux registres, le symbolique et le numérique, néanmoins, aucun de ces deux types de représentations ne facilite la tâche d'identification du sens attribué à l'expression. En lisant l'expression, nous ne pouvons savoir si elle représente la fréquence avec laquelle l'événement A arrive ou si par contre elle symbolise une mesure de la certitude de la véracité de la proposition A. Pour dévoiler son signifié, nous avons besoin d'un travail additionnel dans le registre langagier. Nous situations-problémes s'intéressent précisément aux relations entre ces trois types de registres, le numérique, le symbolique et le langagier, même si nous centrerons nos analyses sur le dernier afin de chercher des traces des interprétations attribuées par les élèves aux différents calculs effectués.

Notre approche se caractérise, entre autres, par une relation hiérarchique entre ces types de registres et le langagier se situe au-dessus du numérique et du symbolique. Avec cette hiérarchie nous souhaitons mettre en évidence une relation où le sémantique détermine le champ du possible par rapport au numérique. En effet, c'est à partir d'un paradigme interprétatif de la probabilité qu'il est possible d'effectuer un ensemble de calculs, en d'autres termes, c'est l'approche interprétative qui détermine l'ensemble de possibilités sur le numérique. La formule de Bayes en est un exemple, c'est en se plaçant dans le paradigme bayésien qu'il est possible de probabiliser sur une hypothèse.

Le lecteur trouvera plus de détail sur la relation entre ces registres avec la formule de Bayes dans le Chapitre I, mais elle est présente tout au long de notre travail. En fait, c'est cette même relation hiérarchique que permet à plusieurs auteurs de repérer des erreurs d'interprétations lors de calculs d'intervalles de confiance ou même de calculs de probabilité (Hacking & Dufour, 2004; Lecoutre, 2005; Régnier & Oriol, 2001). En effet, ces erreurs ne sont repérables que s'il a une correspondance entre l'approche interprétative et le calcul de la probabilité. C'est parce qu'il est impossible que certaines expressions apparaissent dans un paradigme probabiliste donné qu'on repère les erreurs en question. C'est à partir de cette correspondance entre paradigme probabiliste et possibilités de calculs que nous proposons cette relation hiérarchique où la dimension sémantique, plus que le calculatoire, détermine le champ du possible . Notre approche soumet donc le calcul de probabilité à son interprétation, non pas en importance, mais par l'espace de possibilités que le deuxième détermine sur le premier.

Le cadre indéterministe et la dimension sémantique de la probabilité nous ont conduit à nous intéresser un ensemble de concepts et de questions mettant en évidence la singularité du sujet. Pour ce qui concerne le premier ensemble nous pouvons mentionner les notions de qualité de l'information, contextes réels, transparence de simulations ou même les éléments caractéristiques propres à chaque interprétation. Pour ce qui concerne le deuxième, nous nous sommes intéressés par exemple au rôle des contextes réels, tant pour leur possible contribution à interpeller des conceptions déterministes comme pour leur fonction dans la construction de sens d'une interprétation.

La plus part de ces concepts sont proposés dans les deux premiers chapitres et sont finalement intégrés dans le troisième, pour l'analyse de nos expérimentations. Nous avons effectué trois expérimentations, une pour chaque situation-problème. Toutes les trois dans une classe de BTS Electrotechnique de la Région Bretagne. Nous avions souhaité expérimenter dans des plus élémentaires de l'enseignement, bien entendu en testant d'autres situations. Néanmoins et bien qu'ayant trouvé des enseignants motivés, nous avons conclu, après quelques pre-expérimentations, que les contraintes pesant sur le lycée étaient trop nombreuses pour introduire le concept dual de la probabilité. Nous nous sommes donc orientés vers une formation laissant plus de marges de manœuvre qu'une classe de Première ou Terminale. C'est ainsi qu'ayant fait connaissance d'un enseignant de BTS très motivé par le sujet, nous nous sommes décidé à expérimenter au niveau BTS. Néanmoins, rien ne nous semble en principe, au niveau cognitif, pouvoir empêcher d'aborder la dualité de la probabilité à des niveaux plus élémentaires de l'enseignement.

Les trois situations-problémes correspondent aux trois types de contextes retenus, une situation pour le contexte fréquentiste, les deux autres pour les contextes bayésiens. Pour ces problèmes, nous avons considérés deux niveaux de variables. Le premier concerne un ensemble de variables communes aux deux interprétations, le deuxième un ensemble tendant à faire émerger une interprétation tout en cachant l'autre selon les spécificités de chaque problème. Dans le premier niveau, commun aux deux approches de la probabilité, nous nous intéressons aux caractéristiques partagées par les deux approches, parmi elles, nous citons les contextes réels, les prises de décisions et l'encouragement aux débats. Dans le deuxième niveau nous plaçons les variables tendant à favoriser l'émergence d'une notion en cachant l'autre. En effet, la probabilité est un concept se caractérisant par l'imbrication de ces deux interprétations les situations purement fréquentistes ou purement bayésiennes sont rares. Pour maîtriser cette intrication de signifiés nous proposons un ensemble de variables faisant que la situation se corresponde aux éléments caractéristiques propres à chaque approche, de cette manière, en gérant certaines variables nous arrivons à proposer des situations de caractéristiques soit bayésiennes soit fréquentistes. Néanmoins, cette séparation entre approches est fragile et il arrive parfois qu'une situation fréquentiste puisse devenir bayésienne et vice-versa, en modifiant quelques paramètres du problème. Pour décrire se phénomène d'instabilité d'approche dans le contexte d'un problème nous avons introduit le terme d'intrication de signifié.

Etant donnée l'instabilité des situations, la connaissance de la part de l'enseignant des enjeux de la situation devient indispensable, sans cette connaissance lui même pourrait faire basculer involontairement la situation. La nécessité de lui faire connaître l'importance de cette maîtrise nous a conduit à partager avec lui ces idées à travers un document que nous présentons dans cette thèse sous la forme d'une analyse *à priori*. Cette analyse ne répond pas à un détail des tâches nécessaires à la résolution du problème de la part des élèves mais plutôt à un possible déroulement de la séance qui nous permet de préciser un ensemble de questions ouvrant la discussion avec l'enseignant sur les enjeux de la situation. Ces expérimentations seront détaillées dans le chapitre III de cette thèse.

Chapitre I : Enquête épistémologique

1.1 Introduction

La notion de probabilité est un vaste sujet qui a nourri des débats d'ordre philosophique, logique, psychologique et même théologiques le tout en amont de statistiques. Parce qu'elle permet d'aborder rationnellement l'incertain, la probabilité intéresse un grand nombre de sciences qui, en se l'appropriant, l'ont enrichie et rendue plus complexe. C'est ainsi que nous pouvons rencontrer une probabilité logique, une autre propensioniste, une autre bayésienne, etc. Cette complexité nous a suggéré un premier chapitre quelque peu hétéroclite où quelques-uns de ses aspects seront abordés au moins dans leurs principes les plus élémentaires afin de rendre compte de la complexité de l'ensemble. Néanmoins, l'interprétation de la probabilité et son lien avec la statistique inférentielle resteront au cœur de notre travail.

En fait, l'analyse détaillée de cette diversité de la notion de probabilité va nous permettre d'identifier les différences entre les diverses approches de la probabilité en statistique inférentielle et ainsi de déceler les différentes interprétations de la probabilité.

Dans cette première partie, nous présenterons ces différentes facettes de la probabilité et nous développerons des éléments tendant à appuyer notre approche duale de la probabilité. En effet, nous admettrons par la suite que dans un contexte statistique, le terme *probabilité* est associé à deux grands groupes d'interprétations. L'un traditionnellement repéré comme *fréquentiste*, l'autre comme *bayésien*. Si la notion de probabilité est encore influencée par les sciences qui s'en sont servies pour leur développement, ses interprétations possibles se rattachent, par essence, aux deux sources principales que nous venons d'évoquer et c'est pour cela que nous avons utilisé l'idée de groupes d'interprétations.

Pour cela nous prendrons quelques auteurs ayant fortement influencés la construction du concept de probabilité, et nous montrerons qu'il y'en a eu un usage dual. En effet, la probabilité admet deux grands groupes interprétatifs, d'une part elle signifie une attente à long terme, d'autre part un degré de certitude. La première est connue comme *fréquentiste*, la deuxième comme *bayésienne*.

Et nous les définirons bien en tant que groupes parce que ces deux interprétations constituent le "noyau" d'un ensemble d'interprétations auxquels d'autres se rattachent par leur proximité de signifié.

Nous verrons par exemple que la probabilité logique est interprétée comme une mesure de certitude et qu'elle appartient au groupe de la probabilité bayésienne, et que finalement, elle ne sera pas retenue dans notre synthèse comme une interprétation statistique en soi, mais plutôt comme une notion d'un champ scientifique particulier appelé *Logique Inductive* qui partage les principes d'une notion bayésienne (Figure 1).

Interprétations de la probabilité dans les paradigmes inférentiels statistiques

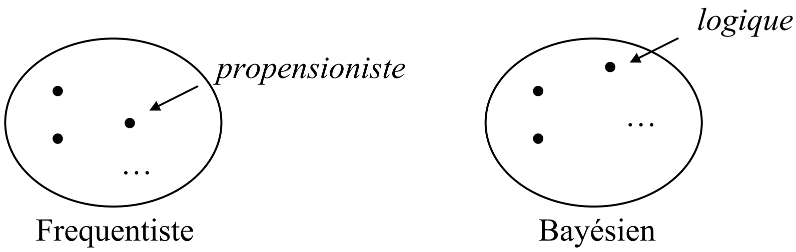


Figure 1

Certains auteurs proposent un spectre beaucoup plus large d'interprétations que celui de la dualité. Quelques-unes de ces autres interprétations (ou définitions) seront évoquées dans ce chapitre, mais toujours présentées dans leur rapport avec la dualité *fréquentiste – bayésienne*. D'autres ne seront que citées car soulignons le : notre objectif majeur reste la probabilité et son rôle en statistique.

Voici donc les deux principaux objectifs pour ce premier chapitre : l'un consiste à repérer des traces de la dualité d'interprétation de la probabilité au long de son histoire, l'autre s'oriente vers une recherche d'éléments caractéristiques à chacun de ces deux groupes d'interprétations.

Ces éléments caractéristiques fonctionneront comme des indices nous facilitant l'identification de l'interprétation de la probabilité dans un problème. Dans une deuxième étape ils serviront pour identifier l'interprétation sous-jacente dans les exercices des manuels (*Chapitre II*), ils seront aussi une référence au moment de concevoir nos situations-problèmes (*Chapitre III*).

D'ailleurs, le lecteur ne trouvera pas ici les critiques versées d'un champ à l'autre, (c'est-à-dire des fréquentistes vers les bayésiens et vice-versa). Notre position n'est pas de trancher sur la validité d'une ou de l'autre de ces approches. En tout cas et en tant que didacticiens, nous nous interrogerons sur l'épistémologie de la probabilité et son évolution dans le processus de transposition didactique (Chevallard, 1985), une probabilité qui étant duale dans la référence épistémologique subit des transformations dans le système éducatif devenant par la suite un objet à enseigner réduit à l'approche fréquentiste, mais dans aucun cas nous ne prendrons position devant ces approches.

Nos premiers chapitres s'adressent donc à montrer que d'une part l'objet épistémologique de référence est dual (*Chapitre I*); et d'autre part, qu'en amont des choix des rédacteurs des documents officiels, il est possible de trouver dans les exercices des manuels des situations tant bayésiennes que fréquentistes (*Chapitre II*). Nous développerons donc nos premières hypothèses et le résumé de notre enquête épistémologique.

1.2 Dualité d'interprétation de la probabilité et recherche d'éléments caractéristiques

Comme nous l'avons signalé, les interprétations de la probabilité ne sont pas toujours faciles à cerner. Pour ce faire, il faut se situer au dessus des calculs et des techniques opératoires. Dans le cas où la probabilité est un objet avec une dimension interprétative, la représentation symbolique ne sera pas de grand secours à l'identification associée à une probabilité donnée.

Par exemple si nous lisons P(A) =1/3, il sera impossible à partir de ces registres (symbolique et numérique) de découvrir à quelle interprétation fait référence l'émetteur de ce message. En effet, les représentations symbolique et numérique ne suffisant pas.

Elles peuvent même entretenir la confusion ; en tant que représentation unifiée pour les deux notions de la probabilité, tout calcul ou symbolisation représentant un lieu d'échange potentiel d'interprétation.

Par exemple : un problème dans lequel la probabilité commence à être interprétée comme *fréquentiste* pourrait, après plusieurs transformations sur le plan calculatoire, "distraitement" devenir bayésienne sans que l'interlocuteur puisse se rendre facilement compte de la faute commise ((Jaynes, 1995) ; (Lecoutre, 2005) ; (Neyman, 1977) ; (Régnier & Oriol, 2001)).

D'une certaine manière, le paragraphe précèdent anticipe notre choix méthodologique pour les chapitres à venir : si dans nos expérimentations il y a des procédés à observer, des erreurs à trouver ou des réussites sur lesquelles compter, ceux-ci ne s'observerons pas facilement en classe de mathématiques, tout simplement parce que les registres favorisés sont en général le symbolique, le numérique et le graphique : aucun d'entre eux ne disposant des outils nécessaires à une interprétation de la probabilité.

Le seul registre capable de le faire est le langagier. Nos observations se focaliseront donc sur les échanges sous les formes orale et écrite effectués en classe. Néanmoins, et nous le verrons plus en détail dans le *Chapitre III* consacré à nos observations, le registre langagier pose aussi ses difficultés. En effet, le mot *probabilité* subit le même sort que son corrélé symbolique P(A), il est utilisé par les deux interprétations ; il ne reste que les expressions synonymiques telles que "chance" et d'autres mots assez vagues pour préciser un signifié si complexe…

Enfin, le fait de travailler sur les interprétations semble s'avérer complexe en classe de mathématiques Nous traiterons plus en détail cette problématique dans les chapitres II et III.

Mais même si le travail d'interprétation de la probabilité en classe de mathématiques est laborieux, il n'est possible que si la situation proposée est cohérente avec l'interprétation projetée pendant la séance. En d'autres termes, si nous souhaitons faire émerger en classe une interprétation en particulier, il faudra créer les conditions contextuelles spécifiques à l'interprétation visée.

Nous nous intéressons précisément dans ce premier chapitre à quelques conditions contextuelles spécifiques à chaque interprétation de la probabilité. Elles seront précisées sous la forme d'éléments caractéristiques associés à chaque interprétation.

Nous proposerons la caractérisation des interprétations de la probabilité par une succession d'approches complémentaires dont chacune de ces approches permettra de comprendre une différence en particulier. Cette caractérisation pourrait être décrite comme celle d'une identification par "couches" de cet objet complexe appelé *probabilité*. La probabilité sera alors vue comme une image assemblée par transparents, où chacun de ces transparents apporte ses pièces à la finalité appelée probabilité.

L'analyse individuelle de chaque couche nous montrera comment les interprétations se différencient sur chaque transparent. Finalement, pour accéder à une vision intégrale de la complexité de la probabilité et de ses interprétations, la solution consisterait à en rassembler les transparents et à les regarder ensemble (Figure 2).

Reperage d'éléments caracteristiques des interpretations de la probabilité

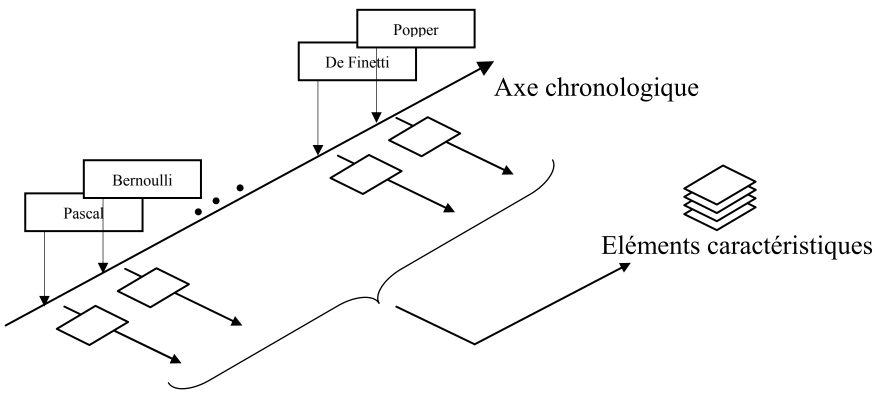


Figure 2

Nous privilégierons un ordre chronologique pour organiser cette recherche d'éléments caractéristiques, ces transparents, pour reprendre la métaphore de l'image composée.

Présentation chronologique qui commencera par les idées de Pascal et terminera par celles de Popper. Nous nous servirons d'exemples historiques pour en retenir des éléments caractéristiques qui seront ensuite repris dans un deuxième et troisième chapitre . Il y aura donc une ligne chronologique depuis laquelle nous tirerons des traces de la dualité de la probabilité et nous croiserons sur cette même ligne le repérage d'éléments caractéristiques.

En d'autres termes, nous chercherons à identifier la présence de la dualité au long de son histoire, et nous présenterons parallèlement les différents éléments caractéristiques des deux interprétations de la probabilité. Dans un premier temps, nous nous contenterons de donner une première approximation de ces deux groupes interprétatifs :

Interprétation fréquentiste

A l'origine de cette interprétation se trouve la présomption de stabilisation de fréquences lorsqu'un certain type d'expérimentation est répétée un nombre infini de fois. La probabilité est alors la limite de la proportion de l'apparition du phénomène en question. Le théorème de Bernoulli (Bernoulli, 1713) est son application mathématique de référence.

Interprétation bayésienne

Cette autre notion a un versant épistémique. Elle exprime une mesure de certitude ou degré de croyance, de crédibilité portée sur une proposition. Le théorème de Bayes (Bayes, 1763) devient pour cette approche, un critère normatif pour reformuler les probabilités en fonction de l'information dont on dispose :l'expression mathématique symbolisant le conditionnant de l'information est la probabilité conditionnelle.

La première de ces interprétations est couramment connue comme *objective* tandis que la bayésienne est associée au terme *subjective* et cela principalement par le conditionnement de l'information lors de l'évaluation. En effet, il est admis dans le modèle bayésien que deux individus puissent évaluer des probabilités différentes sur une même proposition et s'ils disposent d'information significativement différente.

A partir de cette première esquisse de définition des deux interprétations, nous expliciterons notre première hypothèse en la déclinant en un ensemble de propositions plus précises. Nous admettrons par la suite que :

a) la probabilité est une notion relativement récente.
b) elle a émergé au dix-septième siècle dans un contexte d'effervescence d'idées assez particulier favorisant sa configuration duale (Hacking, 2002).
c) dans cette émergence d'idées, plusieurs tentatives de modifications substantielles de cette dualité ont échoué au fil du temps.
d) cette dualité perdure en statistique sous la forme de deux grands courants inférentiels : l'un défini comme inférence fréquentiste, l'autre comme bayésienne.

Le schéma pour ce chapitre

Appuyer ces hypothèses par les faits s'avère difficile. Surtout si l'on souhaite le faire par les registres de l'époque. Au dix-septième siècle, les signifiés des mots n'étaient pas forcément les mêmes qu'aujourd'hui et de plus, les textes mathématiques étaient rédigés en latin. Cela nous a amené à nous référencer sur des travaux d'épistémologues ou de chercheurs ayant consacrés leur étude au sujet qui nous intéresse. Parmi nos principales sources nous reconnaissons l'influence d'Ian Hacking, Glenn Shafer, Gilbert Saporta et Andrew Dale.

A partir de cette première référence épistémologique, nous avons consultés les principaux acteurs de la construction de la notion de probabilité, nous citons entre autres Pascal, Bernoulli, Bayes, Leibniz, Laplace, de Finetti, von Mises, Jaynes, Carnap, Popper et Keynes. La présentation aura l'organisation chronologique précédemment commentée ;depuis la seconde moitié du XVIIème siècle jusqu'à nos jours en nous centrant sur deux grands pôles : l'un le dix-septième et dix-huitième siècle et l'autre le vingt unième siècle.

Parmi les acteurs nous avons pris en compte aussi ceux qui sont considérés comme intransigeants dans le sens d'une position de non reconnaissance de l'"autre" interprétation de la probabilité. Cela à effets de mettre plus en relief les caractéristiques différentielles des interprétations de la probabilité. Notre position déjà exprimée dans nos hypothèses, peut être considérée comme éclectique dans la mesure où nous admettons la cohabitation des deux interprétations en statistique.

Les épistémologues semblent s'accorder sur le début de la notion de probabilité à partir des échanges épistolaires entre Pascal et Fermat. Nous commençons alors par quelques idées de Blaise Pascal.

1.3 Blaise Pascal. Un début de la dualité

Les épistémologues reconnaissent la deuxième moitié du dix-septième siècle comme le moment clef dans la construction de la probabilité. Les lettres entre Pascal et Fermat (1654) autour d'un jeu de partage est un sujet classique en histoire de la probabilité fréquentiste. A cet exemple, nous y en ajouterons un autre qui traite de la probabilité en tant que degré de certitude dans lequel Pascal propose des arguments en faveur de la croyance en Dieu.

Le partage

Il s'agissait d'un problème proposé à Pascal par le Chevalier de Méré où il fallait partager les gains d'un jeu devant être interrompu par des raisons extérieures au jeu même. La solution proposée par Pascal fait intervenir ce qu'aujourd'hui nous appelons les principes de la théorie de la décision, la notion d'espérance mathématique étant au cœur de sa solution.

Soulignons que ni la notion de probabilité ni celle d'espérance étaient d'usage à l'époque, l'attention était orientée vers la recherche d'une structure argumentative de ces démarches de résolution. Tout était en train de commencer dans ce domaine et les objets mathématiques n'étaient que les engrenages d'une machinerie plus vaste de recherche d'une nouvelle rationalité pour traiter l'incertain, où les jeux de dés n'étaient que la pointe visible de cet iceberg ; derrière la résolution de ces jeux naïfs se cachaient des intentions de rationaliser un vaste champ d'actions humaines.

Pour ce qui concerne le problème de partage de Pascal, les ingrédients premiers d'une probabilité fréquentiste étaient présents dans ce problème : jeu de hasard et reproductibilité sous les mêmes conditions sont les plus remarquables. Ce problème constitue une référence classique en histoire de la probabilité, particulièrement en France où il est repris de manière systématique dans les recherches didactiques et sur lequel nous ne reviendrons plus sur cet exemple emblématique de la théorie fréquentiste de la probabilité.

Le pari

Un autre problème de Pascal nous intéresse beaucoup plus à ce moment, il garde les traces d'une probabilité bayésienne et cette première confrontation entre les deux interprétations nous permettra de commencer à trouver des éléments différenciateurs. Il est connu comme *Le pari de Pascal,* probablement écrit en 1658 et apparu dans les *Pensées*[2]

[2] Notre lecture s'est appuyée sur un exemplaire numérisé publié en anglais et distribué par University of California Libraries.

(Pascal, 1670) par les éditeurs de Port-Royal. Pour ce problème, Pascal fait intervenir encore une fois des notions de théorie de la décision, solution qui vient de l'hypothèse qui maximise l'espérance mathématique de l'utilité. Nous ne traiterons pas en détail les hypothèses autour de l'existence de Dieu car c'est en fait la notion de probabilité sous-jacente au problème qui nous intéresse.

Le pari de Pascal se trouve dans un passage intitulé *Infini-rien*. L'épistémologue Ian Hacking décrit le problème de la manière suivante (Hacking, 2002), page 104 et 105:

> *« (...) L'argument s'adresse à un genre de personne qui, non convaincue par les preuves de la religion, et encore moins par les arguments des athées, demeure hésitante entre la foi et incrédulité (...) »*
>
> *« (...) L'une de prémises de l'argument est de dire ou bien il n'y a pas de Dieu, ou bien il y en a un qui est celui dont les caractéristiques sont correctement rapportées par l'Eglise (...) »*
>
> *« (...) Dieu est ou il n'est pas est la tournure exprimant cette partition (...) ».*
>
> *« Le problème décisionnel est formé de deux situations possibles, el de deux actions possibles. Si Dieu n'est pas, les deux lignes d'action reviennent sensiblement au même. Vous allez vivre votre vie sans aucun désagrément consécutif à une intervention surnaturelle. Mais si Dieu existe, parier sur sa non-existence entraîne la damnation. Parier sur son existence peut amener le salut. Le salut est préférable à la damnation. Ainsi l'option Dieu domine l'option Dieu n'est pas. Le problème est résolu par l'argument dominant (...) »*

Nous n'aborderons pas la problématique théologique habillant le problème. Ce qui nous intéresse dans cette argumentation est toujours la notion de probabilité. Notion qui s'avère d'une nature complètement différente de celle du problème de partage du même auteur. Nous verrons que Pascal applique les idées centrales du concept de probabilité bayésienne même si à l'époque ni cette interprétation, ni la fréquentiste ne faisaient partie de ses réflexions.

Mais comment fait-on pour parler de probabilité bayésienne alors que le terme n'est pas présent dans l'essai de Pascal ? Même si le terme était présent, il ne serait pas de grand secours, tout simplement parce que ce mot est employé par les deux interprétations sans distinction. Pour identifier l'interprétation, nous chercherons des éléments caractéristiques à chacune. Ces éléments ou indices apparaîtront dans le contexte et les questions posées par l'interlocuteur.

Les paragraphes qui suivent tenteront d'identifier ces éléments, nous permettant d'infirmer que pour chacun de ces deux problèmes, il y a deux interprétations de probabilité différentes. Nous commencerons par préciser un des éléments nous permettant de contribuer à l'identification des interprétations : l'objet sur lequel porte la probabilité.

Une première différence : Les objets A en P(A)

C'est une des principales différences entre les deux interprétations de la probabilité. Derrière celle-ci se cachent de profondes divergences philosophiques mais nous préférons par souci de clarté et de simplicité, les aborder progressivement et ce à partir de cette différence facilement tangible.

Lorsque nous nous réfèreront à A en P(A) nous entendons couramment utiliser le terme *événement*, c'est ainsi que nous parlons par exemple de la probabilité de l'événement A. Cette terminologie est généralement associée au contexte fréquentiste. Son emploi ici contrarierait nos propos de recherche de différences entre les deux notions. Nous éviterons ce terme qui renvoie implicitement à une de ces notions, nous utiliserons donc le terme *objet* pour nous référer d'une manière neutre à l'argument A en l'application P(A). Un *objet* est ici "quelque chose" sur lequel s'effectue le calcul de probabilité, nous préciserons les types d'objets sur lesquels porte une probabilité.

Les objets A pour la probabilité fréquentiste

La série infinie

Si par définition fréquentiste la probabilité est la limite atteinte par un nombre infini de répétitions du phénomène[3], toutes indépendantes, elle ne peut que faire référence à cette série infinie en question. En d'autres termes, *A* (l'*objet* à probabiliser), doit non seulement être plausible de reproduction, le contexte devant en plus faire référence explicitement à cette attente à long terme.

Dans ce contexte nous inscrivons le problème de partage de Pascal. Bien qu'il est toujours possible de placer l'interprétation bayésienne de la probabilité dans ce jeu[4], notre

[3] En admettant a) toujours la possibilité de réalisation au moins mentalement et b) sous des conditions en principe identiques

[4] Nous faisons référence à l'autre composante du couple « contexte/question posée ». Dans ces paragraphes nous ne traitons que les caractéristiques du contexte.

choix est basé sur les caractéristiques de reproductibilité et de tendance à long terme associées à la probabilité fréquentiste.

Les objets A pour la probabilité bayésienne

Les hypothèses (ou cas uniques)

L'interprétation bayésienne est une mesure, un degré de certitude ou de crédibilité qu'un individu porte sur une proposition. Une des grandes vertus de l'approche bayésienne est que les restrictions sur l'objet A, (proposition A pour une bonne partie de bayésiens) sont minimes voire nulles. L'approche fréquentiste étant empêchée d'aborder des problèmes autres que ceux qui concernent la série infinie, l'approche bayésienne réclame pour elle tous les cas où il faut probabiliser par exemple sur une *hypothèse* ou une *épreuve générique*.

En d'autres termes, tous les cas dont la reproduction du phénomène sur des conditions identiques est non envisageable voire impossible sont par exclusion, des cas bayésiens. Pour cette interprétation, le spectre d'application est beaucoup plus large que pour la fréquentiste, il suffit en principe que nous parlions d'une simple proposition pour que nous puissions lui attribuer une mesure personnelle a notre certitude partielle.

Nous transcrivons deux exemples de son usage courant afin d'illustrer nos propos :

- Sur le site Internet du journal *Le Monde* du 01 Mars 2008 sous le titre « *Le Pentagone préfère EADS à Boeing pour ravitailler ses avions* » nous lisons :

> *«(...) "Je suis extrêmement surprise", a déclaré à l'agence Reuters une analyste reconnue de l'industrie de la défense, Loren Thompson. La probabilité que le groupe européen décroche le contrat était "inférieure à 5 %", selon Yan Derocles, analyste chez Oddo Securities, interrogé par l'AFP (...)».*

- Le Centre de recherches politiques de l'Ecole de Sciences Politique de Paris (Cevipof) publie un rapport intitulé *« Le paradoxe Le Pen (Le Baromètre Politique Français (2006-2007) 3ème vague – Hiver 2006) »*. Sous le sous-titre : *« L'impact du lepénisme: 1 français sur 4 est plus ou moins lepéniste, 1 français sur 2 reste un anti-lepéniste absolu »* nous trouvons une question posée par les chercheurs à un échantillon de la population française (page 1) :

« (...) Est-ce qu'il est tout à fait probable, plutôt probable, plutôt pas probable ou pas probable du tout que vous votiez pour Jean-Marie Le Pen au premier tour de l'élection présidentielle? (...) ».

Dans ces deux extraits, la probabilité porte sur des propositions, des hypothèses pouvant être énoncées respectivement comme *"le groupe européen décroche le contrat"* et *"vous votez Le Pen au premier tour de l'élection présidentielle"*.

Les conditions d'application de la probabilité fréquentiste seraient extrêmement artificielles pour ces deux exemples. Il est difficile d'admettre que l'intention dans ces messages a été de parler d'une attente à long terme. Pour le problème de Pascal (*Le pari*) la situation est équivalente, la notion sous-jacente de probabilité porte sur deux propositions : deux hypothèses disjointes *"Dieu est"* et *"Dieu n'est pas"*, elle est donc bayésienne. Personne ne pouvant s'imaginer dans ce cas, l'existence d'un macro système capable de faire exister Dieu pour le faire disparaître et ensuite le faire exister encore une fois, etc., avec une certaine fréquence d'apparition de x fois … .

Il y a jusque là deux ensembles disjoints d'*objets* sur lesquels nous pouvons probabiliser. D'une part les cas dont la reproduction de l'expérimentation est non seulement envisageable mais aussi évoquée par le problème (interprétation fréquentiste) et d'autre part les cas dans lesquels il se probabilise sur des propositions telles que les *hypothèses* (ou *cas uniques*). Pour ces cas, la reproduction n'est pas envisageable (interprétation bayésienne).

Bien que correcte, la remarque elle, s'avère incomplète. Il y a un troisième groupe d'objets sur lesquels probabiliser que nous appellerons les *épreuves génériques.*

Les épreuves génériques

Les épreuves génériques sont caractérisées par leurs non singularité et leurs possibilités de reproduction, mais, nous ne disposons d'aucune information particulière les concernant. De plus la reproduction de l'expérience qui a généré l'épreuve est toujours plausible à concevoir. Malgré cette possibilité de reproduction, la question ne se pose pas sur la série convergente mais sur une seule de ses réalisations.

Une épreuve générique peut faire partie d'une série infinie ; néanmoins la question, l'objet sur lequel porte la probabilité, n'est pas la série mais une épreuve en particulier de cette série. Un des exemples les plus élémentaires est de s'intéresser au résultat du *prochain* lancer lorsque nous jouons avec un dé. La probabilité fréquentiste strictement parlant ne peut rien apporter à cette question ; elle ne fait que décrire la série et elle ne peut rien dire du prochain tirage. Pour cette interprétation, chaque épreuve n'admet que deux valeurs possibles : l'événement cherché se réalise ou l'événement cherché ne se réalise pas ; la probabilité fréquentiste nous dira quelque chose…Mais à la fin de la série. Ce sera à l'approche bayésienne de donner une réponse à cette question, sujet sur lequel nous reviendrons plus tard (*Chapitre I. Les valeurs logiques pour la probabilité fréquentiste*).

Le problème du pari de Pascal n'est pas du type *épreuve générique*, il n'est pas possible de concevoir qu'il est un parmi d'autres. Il n'est pas possible non plus de concevoir la reproductibilité de l' "épreuve". Nous reviendrons plus tard sur les épreuves génériques lorsque cette synthèse chronologique nous offrira un exemple pour développer ces caractéristiques. Ces cas seront repris constamment dans notre travail et cela pour des raisons diverses :

- Premièrement parce que sa complexité demande plusieurs approches pour mieux la cerner, nous devrons considérer l'approche logique, le type de raisonnement par *référenciation* et en particulier les ensembles de référence au moment de l'évaluation d'une probabilité.
- Deuxièmement parce ce sont des cas très fréquentés en classe et il nous faudra préciser de quelle manière les deux interprétations de la probabilité s'articulent dans ces *épreuves génériques*.
- Troisièmement parce que ces cas nous permettrons d'analyser une interprétation de probabilité appelée propensioniste que Karl Popper (Popper, 1959) et Charles Sender Peirce proposèrent en essayant de donner des réponses à des situations qui échappent au champ d'application de l'interprétation fréquentiste à laquelle ils adhéraient dans un premier temps avec ferveur.

Bref, dans les deux problèmes de Pascal que nous venons d'esquisser, le terme probabilité n'a pas été présent. Néanmoins nous y avons trouvé des traces d'éléments caractéristiques à chacune des deux interprétations, la plus intéressante à nos fins a été le type d'*objet* sur lequel porte la probabilité. C'est ainsi que pour le problème du partage, l'objet est plutôt la *série à long terme* et pour le problème de pari, l'objet est clairement identifié en tant qu'*hypothèse*.

Chronologiquement parlant nous nous intéresserons à un texte contemporain des *Pensées* de Pascal : La *Logique de Port-Royal,* daté de 1662. Ce texte, une référence à l'époque, a été rédigé par Pierre Nicole et Antoine Arnaud. Nous nous centrerons sur les derniers chapitres de la quatrième partie (Chapitres XVI et XVII). Nous nous sommes procuré un exemplaire en français numérisé (Nicole & Arnaud, 1662). Dans ces chapitres nous trouvons les premières traces d'une deuxième étape vers la mathématisation de la probabilité : sa quantification ou évaluation. Nous nous servirons de cet ouvrage pour en tirer de nouveaux éléments caractéristiques.

1.4 La Logique de Port-Royal. Les débuts de la quantification

Des jugements

Les chapitres XV et XVI intitulés respectivement *"Autre remarque sur le sujet de la croyance des événements"* et *"Du jugement que l'on doit faire des accidents futurs"* semblent être les seuls à s'occuper de thèmes liés à la probabilité ((Hacking, 2002), page 115). Ces chapitres contiennent des récits de fait divers (signatures d'actes notariaux, peur de la foudre, etc.) derrière lesquels nous trouvons de manière plus ou moins explicite les idées centrales sur le comportement à adopter rationnellement face à l'incertain.

La probabilité occupe une place très importante dans cette rationalité lorsqu'elle commence à être évaluée numériquement. Nous constatons dès les premières lignes du chapitre XV les principes d'une quantification par un ensemble de référence (cas favorables par rapport aux cas possibles) et des critères élémentaires de décision basés sur la valeur numérique de la probabilité ((Nicole & Arnaud, 1662), page 361) :

> *« (…)Pour juger si on doit les croire ou ne pas les croire, il y en a qu'on peut appeler des circonstances communes, parce qu'elles se rencontrent en beaucoup de faits et qu'elles se trouvent incomparablement plus souvent jointes à la vérité qu'à la fausseté » … « les motifs de croyance qu'il tirait de ces circonstances communes nous avons raison de croire ces événements, sinon certainement, au moins très probablement (…)».*

Nicole et Arnaud, comme Pascal et Fermat, utilisent la probabilité comme un outil rationnel de prise de décision dans des situations incertaines.

Nous décrirons quelques principes d'évaluation en fonction de la nature de l'objet sur lequel porte la probabilité. Un des plus important pour sa récurrence est ce que nous avons appelé l'*ensemble de référence*.

L'ensemble de référence

Dans un contexte de recherche d'une évaluation ou quantification d'une probabilité, nous appellerons *ensemble de référence* une collection d'objets tel que :

- Elle contient l'objet sur lequel porte la probabilité

- Par manque de connaissance sur l'objet en particulier, l'ensemble auquel il appartient devient un repère pour évaluer sa probabilité.

Un exemple banal : il arrive à Grenade un autobus avec 32 touristes dont on sait que parmi eux, 30 portent un appareil photo. On se demande : Quelle est la probabilité que le premier touriste à descendre de l'autocar n'ait pas d'appareil photo ?

L'*ensemble de référence* est ici un ensemble fini composé de 32 touristes dont 2 ont la caractéristique qui nous intéresse (ne pas avoir d'appareil photo). De cette classe, on sait tout (von Mises, 1966) par rapport à la caractéristique en question mais on ne sait rien de cette caractéristique sur un membre de cette classe. Pour nous référer à la caractéristique sur ce membre on fait appel à la classe en tant qu'*ensemble de référence* et la proportion dans la classe devient l'évaluation numérique de la probabilité.

Le procédé d'évaluation d'une probabilité par un ensemble fini de référence a été illustré sur une *épreuve générique* (le premier touriste à descendre du car). Sur le rôle des ensembles de référence,et nous y reviendrons constamment, ils sont en fait un des éléments caractéristiques qui nous facilitent la distinction entre les deux probabilités.

Dans *La Logique* nous commençons à trouver ce genre d'arguments, l'ensemble fini devient une référence où un motif à croire en une proposition. Dans la page 362 où pour croire en la véracité ou la fausseté de la date de signature d'un acte donné et certifiée par des notaires, *La Logique* propose de faire recours à un mode de « *motifs de croyance* » à la proportion d'actes qu'on estime bien datés par rapport à la totalité des actes.

> *« (...) que de mille contrats, il y en a neuf cents quatre-vingt-dix-neuf qui ne le sont (antidatés) que ne le sont point : de sorte qu'il est incomparablement plus probable que ce contrat que je vois est l'un des neuf cents quatre-vingt-dix-neuf et non pas qu'il soit cet unique qui entre mille peut se trouver antidaté(...)»*

Probablement qu'un bayésien se verrait reconnu dans ces arguments. En effet, cette proportion serait une source (et non pas la seule) d'informations lui permettant de quantifier son degré de certitude. Un bayésien pourrait intégrer d'autres informations à l'évaluation d'une probabilité, cette idée semblant être partagée par les auteurs de la *Logique de Port Royal* ((Nicole & Arnaud, 1662), page 362) :

« (…) on y joint d'autres circonstances particulières, comme que ces notaires soient diffamés pour être sans honneur et sans conscience, et qu'ils aient pu avoir un grand intérêt à cette falsification … cela ne me fera pas conclure que ce contrat est antidaté mais diminuera le poids qu'aurait eu sans cela (…) ».

Si l'ensemble de cas favorables parmi les possibles est présenté ici comme une première référence ou comme « *motifs de croyance* », le pas suivant, l'évaluation en soi n'apparaissant qu'à partir du chapitre XVI, pour intégrer la valeur de la probabilité à l'espérance mathématique (ibid., page 366 et 367) :

« (…) Si l'on ne considérait que la perte et le gain en soi, il semblerait que tous y ont de l'avantage ; mais il faut de plus considérer que si chacun peut gagner neuf écus, et n'est au hasard que d'en perdre un, il est aussi neuf fois plus probable, à l'égard de chacun, qu'il perdra son écu et ne gagnera pas les neuf (…)»

Bref, de *La Logique de Port-Royal* nous retenons une évolution en trois directions :

- Premièrement, le concept de probabilité commence à être associé à un mot, le terme en germe est *probabilité*, bien qu'il paye encore ((Nicole & Arnaud, 1662), note de bas de page 361) pour son passé religieux (voir Probabilisme ((Hacking, 2002), page 53-56).
- Deuxièmement, en même temps que les interprétations commencent à se consolider, il apparaît la nécessité d'une quantification, d'une évaluation en termes numériques. Pour *La Logique*, cette mise en nombre sera un outil qui permettra de déterminer le gain espéré, ce dernier étant pris comme facteur décisionnel. C'est le sujet du chapitre XVI où pour décider (et pas seulement en matière de jeux), l'argumentation se base en la maximisation de l'espérance mathématique du gain du jeu.

Troisièmement, les auteurs proposent une source pour quantifier la probabilité (ici bayésienne, certes). Bien qu'assez imprécise, l'ensemble de cas favorables dans les possibles commence à apparaître timidement comme une référence "logique et naturelle" pour le faire. Le schéma de Nicole et Arnaud est assez simple : on s'intéresse à quelque chose qu'on ne connaît pas. On souhaite évaluer numériquement combien y croire. On considère l'ensemble d'objets auquel cette chose appartient

Dans le dernier exemple transcrit de *La Logique*, le contexte est bayésien, l'élément caractéristique différenciateur pris en compte est l'*objet sur lequel porte la probabilité* (un acte en particulier). Pour sa plausibilité de reproduction, l'épreuve est du type *épreuve générique*, et nous parlons de plausibilité de reproduction parce ce qu'il est possible de concevoir un modèle dans lequel l'acte est pris comme un parmi d'autres. Ainsi, vue l'absence d'informations concernant l'objet en question, l'évaluation de la probabilité se fait par un *ensemble fini de référence* : la classe constituée par les actes signés par ces notaires.

Nous venons de décrire très brièvement le procédé d'évaluation d'une épreuve générique par un ensemble fini de référence. Néanmoins, ce moyen n'est pas le seul, un autre ensemble de référence est couramment utilisé pour évaluer une épreuve générique et à la différence du précèdent, celui-ci est infini. En effet, il arrive parfois qu'au lieu de considérer la proportion de cas favorables sur des cas possibles, nous utilisons la fréquence d'apparition du phénomène en question ; ainsi la mesure de sa rareté devient un critère pour évaluer une réalisation singulière, ou épreuve générique.

Les épreuves génériques sont des cas très particuliers et sont aussi un point de convergence des deux interprétations de la probabilité. Lorsque nous nous intéressons à une épreuve générique dont nous connaissons la fréquence d'apparition du phénomène en question, il est habituel et même naturel de combler partiellement le vide d'information sur l'épreuve en particulier par ce qu'arrive fréquemment. De cette manière les deux interprétations sont présentes dans une même situation même si elles remplissent des fonctions complètement différentes, l'une étant un baromètre de nos certitudes (bayésien), l'autre un moyen parmi d'autres de la quantifier (fréquentiste).

Il arrive parfois qu'une épreuve générique puisse être évaluée par les deux types d'ensembles de référence (fini et infini) ; les situations scolaires, nous le verrons, le font fréquemment en rangeant le contexte de manière telle que les deux ensembles donnent la même mesure de certitude. Pour cela, il suffit d'assurer des cas élémentaires équiprobables, cette condition accompagnée par la plausibilité de reproduction qui caractérise les épreuves génériques font que les deux mécanismes d'évaluation soient équivalents en termes d'évaluation numérique. Dans le *Chapitre II*, nous les analyserons plus en détail lorsque nous nous intéresserons aux exercices des manuels scolaires.

Nous introduisons une deuxième différence nous permettant d'avancer dans notre recherche d'éléments caractéristiques différenciateurs des deux interprétations de la probabilité : *La valeur logique en jeu en fonction des objets sur lesquels porte la probabilité.* Cet élément caractéristique nouveau aidera à comprendre le caractère épistémique de la probabilité bayésienne.

Une deuxième différence : la valeur logique des objets A en P(A)

Cette autre caractéristique reste très attachée à la première (*Chapitre I. Une première différence : Les objets A en P(A)*). Néanmoins, cette approche logique qui sera après progressivement approfondie au moment de nous occuper des travaux de Leibniz (Leibniz, 1765) et de Keynes (Keynes, 1921) nous apportera d'éléments différenciateurs dans notre recherche.

Les valeurs logiques pour la probabilité fréquentiste

Imaginons un cas typiquement scolaire : un dé équilibré, ses quatre faces dénommées : w, x, y, z. Nous définissons l'événement x : "la face x est cachée" ; autrement dit : "la face x est en contact avec la table". L'hypothèse d'un dé équilibré nous permet d'en proposer une deuxième, celle de la convergence de la proportion de manifestations de l'événement x vers la valeur ¼ si on lance le dé un nombre infini de fois. Admettons que nous prenons les résultats des dix premières épreuves d'une de ces séries de proportion convergente :

w, w, y, x, y, y, x, z, z, z,...

Pour chacun de ces lancers, il est impossible de prédire son résultat *a priori*, autrement dit, avant le lancer. Un lancer sera un parmi quatre possibilités (x, w, y, z) et on le saura *a posteriori,* après le lancer. Comme nous avons précisé (*Chapitre I. Les épreuves génériques*) la probabilité fréquentiste ne peut pas prédire la valeur d'un tirage en particulier parce qu'elle ne fait que décrire la tendance à long terme.

Dans cet exemple, nous nous intéressons à l'arrivée de x lors du prochain lancer. Chaque lancer prendra la valeur logique "vrai" (si la face x est en contact avec la table) ou "faux" (si la face x n'est pas en contact avec la table), mais cela ne sera valable qu'après avoir effectué le lancer. *A priori*, la probabilité fréquentiste est un outil qui s'avère inadapté à ce genre de questions, car répétons-le, elle ne fait que décrire ce qui arrivera si l'on effectue un nombre infini de tirages, et ne nous dit rien du prochain lancer.

En résumé, pour la probabilité fréquentiste, chaque épreuve assume deux valeurs logiques possibles (vrai ou faux) et toujours *a posteriori* de l'épreuve. Pour ce qui concerne la série, sa valeur logique est assumée comme vraie par définition fréquentiste.

Les valeurs logiques pour la probabilité bayésienne

La notion de probabilité bayésienne en tant que degré de certitude est une mesure de "croyance" : plus sûr on est de la véracité d'une proposition, plus proche la probabilité sera de 1. Les valeurs 1 et 0 sont deux valeurs de certitude, le vrai et le faux, à savoir la certitude de la véracité et la certitude de la fausseté. Parmi ces deux valeurs logiques extrêmes, toute une gamme continue de valeurs logiques intermédiaires est possible.

Reprenons le problème du contrat signé par les notaires ((Nicole & Arnaud, 1662), page 361). Cet exemple se trouve dans le chapitre XV consacré à établir des critères tendant à décider sur la croyance des événements passés. Nous soulignerons l'importance de cette variable chronologique "passé", elle nous facilitera la tâche de mise en évidence des différences en minimisant les intersections entre les deux interprétations[5].

Imaginons donc un certain Pierre Dupont dans une situation proche à celle décrite dans *La Logique* : Monsieur Dupont a un acte sur son bureau. Cet acte a été signé par deux notaires (admettons il y a une semaine). Monsieur Dupont s'interroge sur la possible falsification du document.

Entre M. Dupont et son acte signé, il y a un rapport épistémique, relatif à sa propre connaissance. Monsieur Dupont ne connaît pas la vérité sur l'authenticité ou non de son acte.

Monsieur Dupont cherche alors à évaluer la situation en termes de degrés auxquels il serait raisonnable de croire en la vérité de la proposition : *l'acte n'est pas antidaté*. En cherchant des raisons, un motif, une information pertinente lui permettant de trancher, il ne trouve rien en particulier. M. Dupont remonte donc à l'ensemble des actes auquel le sien appartient et il apprend que très peu d'actes sont antidatés. En autres termes, M. Dupont suit les indications de *La Logique* (de Port-Royal) : il apprend que sur mille actes, un seul est antidaté. Il utilise cette information pour trancher sur son acte : "il est fortement probable que mon acte ne soit pas antidaté".

[5] Cette terminologie a l'intention d'anticiper sa fonction didactique (variable didactique) que nous traiterons dans le chapitre III

Nous souhaitons bien préciser le rôle de cet ensemble de référence dans une démarche de quantification de la probabilité pour une épreuve générique. Cet ensemble n'est pas le seul moyen à être utilisé pour évaluer numériquement une probabilité. Même *La Logique* nous en propose d'autres : toute information considérée comme significative (telle qu'en savoir plus que ces deux notaires) pouvant faire basculer le degré de croyance.

Le cas de l'acte de M. Dupont nous suggère un contexte bayésien. Les enjeux amènent à penser que Monsieur Dupont[6] est intéressé à son acte en particulier même s'il cherche ailleurs pour décider. Cette probabilité de M. Dupont est une mesure de certitude, mesure qui prend une valeur quelque part entre deux certitudes et dont la proportion d'actes antidatés n'est qu'une information donnant des motifs lors de l'évaluation.

Si l'on analyse le cas de M. Dupont du point de vue de notre première caractérisation (les objets sur lesquels s'opère la probabilité) on observe que celui-ci est une *épreuve générique*, principalement par sa plausibilité de reproduction. Une analyse du point de vue logique nous indique qu'elle est une valeur comprise entre zéro et un, et que l'ensemble d'actes signés intervient comme un ensemble de repérage ou de référence, un motif de croire.

Pour l'instant nous disposons de deux éléments caractéristiques complémentaires ou deux transparents : les types d'objets et leurs valeurs logiques. D'ailleurs, l'exemple des actes nous a permis de préciser quelques éléments du procédé de quantification, plus précisément le rôle des ensembles de référence dans l'évaluation d'une épreuve générique.

En continuité avec ces idées et en creusant encore plus l'approche bayésienne, nous trouvons les travaux de Leibniz. Les idées de preuves, de motifs à croire lui intéressaient, et cela pour des raisons de jurisprudence ((Leibniz, 1765), pages 529, 538, 539). Nous consacrerons quelques lignes à ses travaux, principalement à son approche de la probabilité du point de vue logique.

[6] Probablement un « Bureau de l'administration nationale des fraudes » puisse s'intéresser strictement à la fréquence d'apparition d'actes antidatés, mais pas certainement M. Dupont.

1.5 Gottfried Leibniz. Vers une probabilité logique

Parmi ses publications, nous retiendrons l'œuvre intitulée *New essays concerning human understanding* (Leibniz, 1765), ce texte fut achevé en 1705 mais il n'est apparu qu'en 1765. L'organisation sous la forme d'un dialogue imaginaire entre deux personnages appelés Philalèthe et Théophile lui a permis de répondre à la publication de Locke *Essais sur l'entendement humain.* Dans *New essays,* Philalèthe représente les idées de Locke et Théophile celles de Leibniz.

Le discours de Leibniz, bien que parfois flou, est plus analytique et abstrait que celui de la *Logique de Port-Royal*. Il tente de développer des concepts en cherchant un nouveau cadre logique, une logique qui systématise et rationalise *« our belief »* (ibid., page 538). Dans cet ouvrage l'auteur manifeste son intérêt pour une probabilité au delà des jeux de hasard «(…) *the degrees of probability, we still lack that part of Logic* (…)» (ibid., page 213).

Ce texte est particulièrement intéressant pour l'attention que l'auteur porte sur ce qu'à l'époque on appelait les aspects philosophiques de la probabilité. Dans cet ouvrage Leibniz cherche à expliquer l'importance de trouver des éléments tendant à valider certains procédés argumentatifs non déductifs ; pour cela il introduit la probabilité comme une mesure de crédibilité. Dans ce livre, Leibniz s'intéresse plutôt aux aspects sémantiques de la probabilité qu'aux calculatoires. Ces aspects sémantiques sont indispensables à la formulation d'une Logique pour rationaliser l'incertain, nous analyserons son approche interprétative de la probabilité.

Des contributions sémantiques

La logique qui intéresse Leibniz est basée sur une probabilité représentant un degré de certitude qui, en se différenciant clairement du probabilisme (ibid. page 419), soit conforme au constat des observations. Ian Hacking signale sa contribution conceptuelle ((Hacking, 2002), page 132):

> *«(…) Il n'apporta pas de contribution aux mathématiques de la probabilité, mais la conceptualisation qu'il en fit eut de fait un impact durable. La plupart de ses contemporains partaient de phénomènes aléatoires –les jeux de hasard, ou bien la mortalité- et, par un saut imaginatif, spéculaient sur un transfert de la doctrine de chances à d'autres cas d'inférence en contexte d'incertitude. Leibniz prit la probabilité numérique comme notion épistémique première. Les*

degrés de probabilité sont des degrés de certitude. Ainsi il considérait que la doctrine de chances ne concerne pas les caractéristiques physiques des configurations des jeux, mais a trait à notre connaissance de ces configurations (…) ».

En effet, en comparant la probabilité épistémique avec la certitude, Leibniz y voit une nouvelle sorte de preuve (ibid, page 420):

«(…) [I] think that to these species of certitude or certain knowledge you can add the knowledge of the probable; thus there will be two sorts of knowledge as there are two sorts of proofs, the first of which produce certitude, and the second end only in probability(…)»

Il prenait comme contexte son expérience dans le champ judiciaire. L'attestation d'un témoin ne pouvait pas l'amener à la certitude, mais il affirmait que cette information ne devait en aucun cas être écartée, que quelque part ces témoins appuyaient la thèse en confirmant ou pas la véracité des faits. Son intérêt était de systématiser ce procédé. Pour cela il lui fallait une structure logique qui fonde les principes d'une telle démarche. Dans ce contexte la probabilité comblait une place vide permettant de modéliser la crédibilité.

Dans sa recherche pour établir avec précision dans quelle mesure une information pourrait être intégrée à la décision d'un jugement, il considérait que la probabilité était l'outil approprié. Même s'il lui manquait les formalismes appropriés il insistera sur les principes sémantiques d'une probabilité conditionnelle dans laquelle le conditionnant est une information attestée et l'argument sur lequel se probabilise une proposition, à évaluer. Son intérêt pour le rôle de témoins et d'évidences factuelles est reçurent tout au long de son ouvrage (ibid., pages 529, 538 et 539, (Hacking, 2002), page 133).

Leibniz est devenu une référence pour ceux qui au XXème siècle ont tenté de construire une théorie logique. Carnap (Carnap, 1947) et Keynes (Keynes, 1921) ont approfondit cette ligne de recherche, que nous ne ferons qu'évoquer (opportunément), vu notre intérêt pour les enjeux de la probabilité en statistique.

D'après Ian Hacking ((Hacking, 2002), page 173) c'est Leibniz même qui introduit l'usage de la formule de cas favorables et de cas possibles basée sur l'équipossibilité :

« (…) On croit souvent que l'origine de ce concept se trouve chez Laplace, aux alentours de la fin du dix-huitième siècle, alors que c'était un lieu commun dès

l'origine. Il est vrai que Laplace définit la probabilité comme le rapport du nombre de cas favorables sur celui du nombre total des cas également possibles, mais Leibniz faisait de même en 1678 (…) ».

Pour l'instant nous souhaitons retenir deux idées des textes de Leibniz: l'une est son intérêt en une nouvelle sorte de logique pour traiter l'incertain, et qui s'appuie sur l'interprétation bayésienne de la probabilité, l'autre concerne les critères d'évaluation numérique de la probabilité, en particulier ceux qui se basent sur les ensembles de référence pour une épreuve générique.

Une démarche d'évaluation qui à l'époque était justifiée par sa naturalité et qui est devenue progressivement un automatisme. Un de ces ensembles de références est utilisé dans la formule de cas favorables et de cas possibles, l'autre nous le verrons sera la fréquence d'apparition.

1.6 Jacques Bernoulli. Le théorème limite

L'œuvre la plus célèbre de Bernoulli concernant la probabilité reste son livre intitulé *Ars Conjectandi* (Bernoulli, 1713), ce texte posthume, paru grâce aux diligences de son neveu qui finalement le porte chez l'éditeur.

Bien qu'il soit possible de télécharger quelques morceaux numérisés sur des sites de bibliothèques françaises, la quatrième partie, où Bernoulli développe sont théorème et s'exprime en termes de probabilité, ne se trouvait pas disponible sous format numérisé au moment dont nous effectuions la recherche, nous nous sommes procuré donc un exemplaire en anglais sur support papier.

Cet ouvrage comporte quatre parties, en la première partie, l'auteur s'intéresse aux jeux de hasard, en la deuxième, aux combinaisons et en la troisième il propose une suite d'exercices. C'est en la quatrième partie d'*Ars Conjectandi* que Bernoulli se consacre à la probabilité et à son théorème sur la convergence de la série, cette partie (*« The use and applications of the preceding donctring in civil, moral and economics matters »),* comporte cinq chapitres.

Nous pourrions résumer cette cinquième partie comme une étant contribution principalement sur le plan mathématique. En effet, c'est là qu'est démontré le théorème fondamental de la limite. Sur ce théorème, l'épistémologue Ian Hacking (Hacking, 2002) explique l'approche de Bernoulli :

> *« Il part (...) d'in dispositif aléatoire permettant de réitérer des essaies. Il existe une chance et inconnue, p, de succès à une épreuve quelconque. Après n essaies, on relève une proportion s_n de succès. Bernoulli prouve ce que l'on appelle maintenant la loi faible des grands nombres : la probabilité d'une série de n termes pour laquelle $|p - s_n| < \varepsilon$ tends vers 1 lorsque n tend vers l'infini. De plus pour toute erreur donnée ε, il montre comment calculer un nombre n tel que la probabilité que s_n se trouve dans l'intervalle [p- ε ; p+ε] dépasse elle-même toute probabilité donnée $1-\delta$ ».*

Ce théorème constitue un pilier de la théorie de la probabilité, il se trouve dans un grand nombre d'ouvrages sur probabilité, nous n'en rentrerons pas en détails. Néanmoins, ce que semble assez peut repris de Bernoulli est sont approche sur la probabilité. Il commence la quatrième partie en précisant qu'il aurait, deux sortes de certitudes, une objective, l'autre subjective. A l'époque, la première signifiait *« décidé par les dieux »* (Hacking, 2002) ; la deuxième, la subjective, représenterait *«the mesure of our knowledge concerning this truth »* (Bernoulli, 1713). En suite, il précise ce qu'il entend par probabilité (ibid, page 315):

> *« probability is degree of certainty, and differs from the latter as a part differs from the whole »*

Il n'y a dans l'interprétation de cette définition, des doutes d'un possible anachronisme, toutes les sources épistémologiques nous l'ont confirmé, Bernoulli faisait du terme probabilité, un usage dual (Dale, 1999; Hacking, 1971; Shafer, 1996), même la traductrice de l'ouvrage, Edith Dudley Sylla, souligne dans ses notes finales l'approche épistémique d' l'auteur (Bernoulli, 1713) :

> *« When the word probability is definied at the begining of Part IV, therefore, it introduce something new into the mix, namely a mathematics of probability in its epistemic sens (...) one thing ...is called more probable than another if it has a larger part of certainty (...)»*

Pour ce qui concerne l'approche fréquentiste chez Bernoulli, cela peut être trouvé sans difficulté dans son travail sur la théorème limite, nous ne citerons pas ici ce qu'est largement répandu de l'auteur. Il y a donc dans les pensées de Bernoulli une alternation de signifiés, dans quelques paragraphes le terme probabilité signifie un degré de certitude, dans d'autres, la fréquence d'apparition d'un phénomène.

Du point de vue des aspects interprétatifs de la probabilité donc, nous confirmons chez Bernoulli, une dualité de signifié. Pour ce qui concerne les aspects calculatoires, nous observons la reprise de quelques exemples simples déjà abordés en *La logique de Port Royal* (Nicole & Arnaud, 1662). En effet, Bernoulli, et de la même manière que Nicole et Arnaud, a cherché à préciser des critères rationnels pour agir devant l'incertain. Dans ce projet, la probabilité occupe pour Bernoulli, une place dominante (ibid, page 317):

> *« To conjecture about something is to mesure its probability. Therefore we defined the art of conjecture, or stochastics, as the art of measuring the probabilities of things as exactly as possible, to the end that, in our judgements and actions, we may always choose or follow that which has been found to be better, more satisfactory, safer, or more carefully considered ».*

Dans son intérêt pour créer une doctrine pour agir rationnellement en affaires si divers (civiques, morales, et économiques), il considère des situations variées, quelques-unes abordables à partir d'une approche fréquentiste, d'autres, à partir de l'approche bayésienne. Pour ces dernières, la nature épistémique du problème ne laisse pas de doute, il s'agit pour la probabilité de mesurer un degré de certitude de la véracité d'une proposition. Par exemple, Bernoulli s'intéresse à analyser de quelle manière peuvent interagir différents types d'arguments pour déterminer si Maevius a tué Titius (ibid, page 318), ou tel que l'ont fait les auteurs de *La logique de Port Royal*, si un acte est antidaté ou non.

Pour traiter ce genre d'affaires, Bernoulli propose de se baser en une mesure appelée probabilité. Pour l'évaluation ou quantification de cette mesure il propose deux grands moyens et les présente à travers deux exemples, chacun renvoyant à des critères différents d'évaluation. Le premier basé sur la proportion de cas favorables sur possibles, l'autre sur la fréquence d'apparition de phénomènes similaires.

Pour le premier, celui de la proportion de cas favorables et possibles, Bernoulli reprend l'exemple de *La logique de Port Royal*, celui de l'acte d'authenticité douteuse. Il propose une évaluation basée en la proportion d'actes antidatés sur le total d'actes possibles (ibid page 322):

> *« It's clear from the foregoing that any argument's power of proof depends upon the number of cases in which the argument can exist or not exist, indicate or not indicate or event indicate the contrary (...)».*

Nous appellerons ce raisonnement référenciation par des ensembles fini, en effet, pour évaluer combien croire en la fausseté d'un acte en particulier, la méthode consiste en chercher une référence, celle de l'ensemble auquel l'acte appartient, dans ce cas la référence est un ensemble fini (voir *La référenciation*).

L'autre critère d'évaluation est couramment connu comme principe fréquentiste, la question étant toujours épistémique, le critère d'évaluation d'une épreuve se base, cette fois-ci, sur la fréquence d'apparition du phénomène en question, dans ce cas se produit aussi une référenciation, mais à la différence du critère précédent, elle serait par un ensemble infini (ibid, page 327) :

> *«If for exemple, there once existed three hundred people of the same age and body type as Titius now has, and you observed that two hundred of them died before the end of the decade, while the rest lived longer, you could safely enough conclude that there are twice as many cases in which Titius also may*

die whithin a decade as there are cases in which he may live beyond a decade(...)».

Bernoulli, ne s'adresse pas en termes de référenciation, pour lui et dans cet ouvrage, ce critère n'est qu'un exemple pour introduire une notion centrale dans son théorème, elle concerne l'importance de la taille de l'échantillon dans l'étude de la convergence de la série. C'est nous, qui dans notre intérêt de déceler les approches de chacun des auteurs, retenons et caractérisons les différents concepts retrouvés dans les différents textes.

D'une manière très synthétique, nous venons de repérer l'usage dual de la probabilité dans l'ouvrage de Jacques Bernoulli, pour plus de détails des enjeux de cette dualité et de sa constante utilisation dans son théorème, nous renvoyons au lecteur au chapitre dix sept de l'ouvrage intitulé *L'émergence de la probabilité* (Hacking, 2002). ous continuerons avec un auteur, dont sa principale contribution, aussi sur le plan mathématique, est célèbre pour semer les bases de l'inférence bayésienne.

1.7 Thomas Bayes. L'inversion de la probabilité

Les écrits de Bayes ont été publiés dans *Philosophical Transaction* l'année 1764. Oeuvre posthume, c'est Richard Price qui a trouvé l'essai (Bayes, 1763) parmi d'autres documents de Bayes et le trouvant digne d'intérêt l'a envoyé à John Cantor. Le texte de Bayes proprement dit est assez concis et decontextualisé. Les définitions précèdent le développement de la démonstration de son fameux théorème qui reste sur un plan mathématique.

Bien qu'il consacre quelques lignes aux termes *chances* et *probabilité*, sa définition de la probabilité est *opératoire* dans le sens où elle traite des aspects algorithmiques de la probabilité et non pas des aspects *sémantiques* (ibid., page 4) :

> *«(...)Probability of any event is the radio between the value at which an expectation depending on the happening of the event ought to be computed, and the chance of thing expected upon it's happening(...) ».*

Par sa décontextualisation, il est difficile, voire impossible de trouver les traces des interprétations de la probabilité dans le texte original de Bayes. En fait c'est Price qui joindra des exemples ((Todhunter, 1865), page 295) et qui à la fin de l'article expliquera des possibles applications des problèmes de cause à effet. En d'autres termes, des problèmes dont l'observation d'un effet servirait à confirmer l'hypothèse qui rend l'observation plus probable.

Mais Bayes ne fait jamais allusion à ce genre de raisonnements ; dans son essai présenté sous forme mathématique, on ne trouve pas d'applications aussi importantes que celles de Price. Et de plus, il ne présentera pas son théorème pour un ensemble discrets d'hypothèses. Il habille sa démonstration avec un exemple de lancers successifs de deux boules sur une table de billard carrée de telle façon que les lois de probabilités de leurs points d'arrêt sur le tapis soient uniformes ((Droesbeke, Fine, & Saporta, 2002), page 8) :

> *« (...) Sur la* [Figure 3] *(...) ABCD représente le billard carré. Une boule W est lancée dans un premier temps et, à son immobilisation, une droite verticale os est tracée, passant par le point d'arrêt. Cette droite défini deux rectangles CBos et DAos situées à gauche et à droite de os. Une deuxième balle O est lancée : on parle de succès si O s'arrête dans CBos et d'échec dans le cas contraire. On obtient ainsi une séquence de deux événements pour lesquels on peut déterminer des probabilités conjointes et des probabilités conditionnelles, grâce à une approche géométrique, selon le procédé suive par Bayes(...) »*

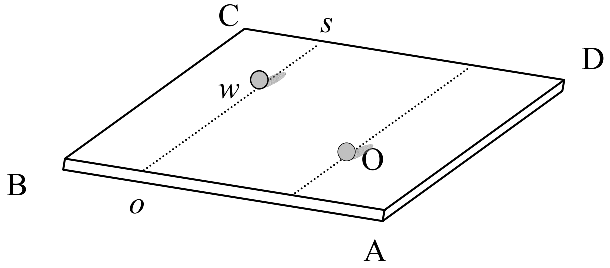


Figure 3

Mais ce degré d'abstraction ne veut pas dire que les enjeux d'un tel problème lui échappaient. Des indices tels que démontrer par des voies différentes les probabilités $P_A(B)$ et $P_B(A)$ et des expressions telles que *"ma chance d'avoir raison"* et *"une seule épreuve"* font penser que Bayes comprenait les implications d'un tel procédé.

Des contributions opératoires

A la différence de Leibniz dont sa contribution a été plutôt conceptuelle, celle de Bayes est mathématique : Il a trouvé une expression permettant d'inverser le schéma du problème de Bernoulli. Comme le dit Price et tant d'autres, l'expression permet de traiter mathématiquement des situations où il faut expliquer un phénomène donné par la recherche d'une cause, la plus probable. A cet effet Jordan affirmait ((Jordan, 1926), page 121):

> *« ...Pour résoudre ce problème, on a recours au théorème de Bayes. Un événement peut avoir pour cause C_1, C_2, C_3,...Désignons par w_i la probabilité a priori, c'est-à-dire avant l'arrivée de l'événement, pour que la cause C_i soit en jeu ; de plus désignons par p_i la probabilité pour que l'événement se produise si la cause C_i agit. On démontre que la probabilité a posteriori, c'est-à-dire après l'arrivée de l'événement, pour que celui-ci soit dû à la cause C_i est*
>
> $$P_i = \frac{w_i p_i}{\sum w_s p_s} \text{ »}$$

L'inversion de la probabilité n'est pas un fait anecdotique, il ne s'agit pas d'un simple échange de places. Nous allons voir qu'elle met plus en évidence encore les différences entre les deux interprétations de la probabilité et que pour pouvoir appliquer cette inversion, on est amené à interpréter la probabilité comme étant un degré de certitude.

Quelle probabilité pour la formule de Bayes ?

La probabilité fréquentiste s'applique dans une seule direction (Batanero, Henry, & Parzysz, 2005): des hypothèses vers la série infinie : en admettant les hypothèses (*H*) il est possible de déterminer la limite de la série ($P(B)$). En général, ces hypothèses (par exemple H : le dé est équilibré) restent implicites dans la symbolisation fréquentiste, son écriture complète prend la forme $P_H(B)$ (en admettant le dé équilibré la proportion d'apparition à long terme de l'événement B est…).

La probabilité fréquentiste ne peut s'appliquer à l'expression inverse $P_B(H)$, même si algébriquement il n'y a aucune contrainte. En effet, pour ce paradigme, l'inversion devient impossible à cause des caractéristiques de l'objet *H*, qui en tant qu'hypothèse, même si inconnue, reste fixe. C'est la différence entre *variable aléatoire* et *paramètre inconnu* ; la variable aléatoire est caractérisée par une incertitude de type fréquentiste, dans lequel la répétition de l'expérience produit des résultats différents ; le paramètre inconnu renvoie à un autre type d'incertitude, de type épistémique.

Pour appliquer la probabilité fréquentiste à une hypothèse, on devrait pouvoir concevoir une expérience dans laquelle on obtiendrait quelquefois l'expression "le dé est équilibré", d'autre fois *"le dé n'est pas équilibré"* et ainsi en reproduisant l'expérience jusqu'à l'infini, la fréquence d'apparition de l'événement "le dé est équilibré" convergerait vers une valeur…Une telle expérience, est bien évidement impossible.

A l'origine de cette faute se trouve la dualité de la probabilité, et elle n'est pas identifiable sur le plan mathématique mais sur le plan philosophique de ces deux paradigmes. Cette faute met en évidence l'importance de l'interprétation de la probabilité dans les manipulations numériques et symboliques.

La probabilité bayésienne, et à la différence de la fréquentiste, admet la possibilité de probabiliser sur une hypothèse *H,* nous l'avons signalé dans notre premier élément caractéristique : l'objet sur lequel porte la probabilité. En tant que degré de certitude, elle peut être appliquée à une hypothèse car *H* n'est qu'une proposition.

A l'époque de Bayes, bien qu'il y a avait des confusions entre ces deux types de probabilités conditionnelles ((Condorcet, 1805), page 65), un grand nombre de mathématiciens reconnaissaient déjà les implications d'une telle démarche. Price s'est aperçu du type de problème auquel la découverte de Bayes pouvait contribuer et il l'a signalé à la fin de l'essai. Les deux probabilités coexistaient depuis longtemps, et leur quantification commençait à se consolider, mais il manquait une expression permettant d'aborder le problème de quantifier combien les faits (*B*) confirmeraient une hypothèse (*H*). C'est grâce à ce théorème que l'on peut mener sur le plan mathématique des questions telles que : ayant constaté un symptôme (*B*) quel diagnostique établir (*H*) sur un patient? Quelle maladie (*H*) est la plus probable vis-à-vis cette information (*B*)?[7]

Pour avancer sur ce genre de problèmes, la probabilité bayésienne s'avère idéale (au moins par rapport à la fréquentiste) ; néanmoins il a fallu attendre plus d'un siècle pour que sa théorie inférentielle se consolide et construise ses objets de base ; certes le même temps qu'a prit la théorie fréquentiste à se développer.

Nous venons de montrer comment la formule de Bayes requiert l'interprétation de la probabilité comme un degré de certitude pour qu'elle ait du sens, pour le faire, nous nous sommes servis du type d'objet sur lequel porte la probabilité.

Nous avons présenté des champs d'applications possibles pour les probabilités conditionnelles pour chaque paradigme probabiliste, et cela en fonction des types d'objet que chaque expression admet. Nous continuerons en introduisant une troisième différence entre les interprétations de la probabilité : *le type de raisonnement mobilisé.*

Nous avons présenté deux types de différences déjà, la première concerne l'*objet A en P(A)* et la deuxième les *valeurs logiques* que chaque interprétation mobilise. Le schéma suivant résume ces deux premières différences (Tableau 1).

Deux des différences entre les interprétations de la probabilité

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="3">Interprétations</td></tr>
<tr><td>Fréquentiste</td><td colspan="2">Bayésienne</td></tr>
<tr><td rowspan="2">Différences</td><td>Objet sur le quel probabiliser</td><td>Série infinie</td><td>Epreuve générique</td><td>Hypothèse</td></tr>
<tr><td>Valeurs logiques portés sur ces objets</td><td>Vrai- faux</td><td colspan="2">Intervalle continue [0,1]</td></tr>
</table>

Tableau 1

[7] D'autres exemples à l'époque ont proposés par Laplace dans sa *Mémoire sur la probabilité des causes par des événements* ((Laplace, 1771)) où l'auteur accompagne avec un sens pédagogique le développement des idées sur la probabilité d'hypothèses.

Avant de décrire les *types de raisonnements mobilisés* pour chaque type d'interprétation nous préciserons en quelques lignes certains enjeux des *épreuves génériques* (probabilité bayésienne). Les *hypothèses* sont des objets plus simples à cerner que les épreuves génériques car ils concernent généralement les suppositions ou conjectures de base d'un modèle statistique. Les *épreuves génériques* mettent en place des démarches assez particulières que nous décrirons par la suite.

Quelques précisions sur les épreuves génériques

Probabiliser sur une épreuve générique veut dire que la probabilité porte sur un objet perçu comme non singulier et dont la répétition de l'expérience est plausible. Pour ce genre d'objets que nous désignons comme génériques, l'évaluation numérique se fait par deux types d'ensembles de référence :

- L'ensemble infini donné par la fréquence d'apparition du phénomène lors de la répétition de l'expérience.
- L'ensemble fini déterminé par les cas favorables dans les possibles.

Nous proposons un exemple pour chacune de ces deux types d'évaluation :

- Je "sais" que dans 30% des cas ma punaise tombe sur tête. Ma punaise a été lancée. Quelle est la probabilité qu'elle retombe sur tête ?.
- Le prochain examen fait peur à Camille (C), son père (P) en essayant de la rassurer lui dit :

  P : *« Ne t'inquiètes pas, 80% des élèves réussissent l'examen ».*

  C : *« mais non c'est pas comme çà pour nous, je viens de la banlieue ».*

Camille semble vouloir dire qu'elle se sent part d'un autre ensemble de référence, plus cerné que le total d'étudiants passant l'examen. Pour elle, son ensemble de référence semble être celui des adolescentes de banlieue (qui passent l'examen) dont le taux de réussite est sensiblement inférieur. Il y a ici un conflit d'ensemble de référence, l'un général : la totalité d'étudiants, l'autre plus restreint : celui des candidats provenant des alentours de Paris.

Ce conflit d'ensembles de référence est en fait un "conflit de recherche des raisons", de recherche de motifs à croire ((de Finetti, 1937), page 21), nous reviendrons sur ce sujet si cher à la probabilité bayésienne.

Une épreuve générique est donc une réalisation sur laquelle nous n'avons aucune information particulière et en ayant besoin de quantifier sa probabilité, nous faisons recours à des ensembles dont elle fait partie, des ensembles tels que la fréquence d'apparition ou la proportion de cas favorables dans les cas possibles.

En fait, la *Logique de Port-Royal* nous a offert un exemple de recherche d'ensemble de référence. Dans le cas de la signature des deux notaires sur un acte, Nicole et Arnaud nous ont proposés de chercher des motifs à croire dans un ensemble fini de référence :

> *« (...) que de mille contrats, il y en ait neuf cents quatre-vingt-dix-neuf qui ne le soit (authentiques ?)(...) »*, et en suite ils nous ont suggéré d'intégrer à l'évaluation d'autres informations: *« (...) on y joint d'autres circonstances particulières, comme le fait que ces notaires soient diffamés pour être sans honneur(...) ».*

Nous avons décrit les épreuves génériques comme des cas bayésiens en tant que plausibles de reproduction et caractérisés par un manque d'information particulière, elles sont menées à être quantifiées par des ensembles de référence Nous présenterons tout de suite les types de raisonnements mobilisés pour chacune des interprétations de la probabilité.

Nous entendons par raisonnements des formes de procédés délibérés ((Chiasson, 2005), page 2) plus ou moins acceptées et "reconnues" à une époque munie de principes et de critères propres, permettant la validation de propositions au sein de son paradigme (Crombie, 1980), c'est ainsi que par exemple le terme "valider" devient plus large que celui adopté en mathématiques dont les valeurs logiques n'admettent que la démonstration ou la réfutation.

Une troisième différence : les types de raisonnement mobilisés

Pour la probabilité fréquentiste

La déduction

Nous avons déjà commenté que la probabilité fréquentiste s'applique dans une seule direction : des hypothèses vers la série infinie. Nous pourrions décrire ce raisonnement comme du type déductif, dan le sens où c'est une procédure par laquelle on conclut à partir de la prémisse admise (hypothèse) vers la série qui en résulte. Une cohérence est à observer : ce type de raisonnement est en consonance avec les valeurs logiques que traite cette probabilité.

Une définition de la déduction dans les termes "allant du général au particulier" ne doit pas nous faire penser que pour la probabilité fréquentiste, on part des hypothèses à un élément de la série (une épreuve générique en fait), cette présomption est invalidée par la valeur logique qui prend la réalisation d'un élément de la série (entre le vrai et le faux) et les valeurs logiques possibles dans une démarche de déduction. Pour la probabilité bayésienne, deux types de raisonnements seront précisés : l'*abduction* et la *référenciation*.

Pour la probabilité bayésienne

L'abduction

Charles Sender Peirce dans ses recherches sur le pragmatisme s'est intéressé au sujet des inférences et leurs critères de validation (au sens large, encore une fois). C'est ainsi qu'il a proposé le terme *abduction* pour un type d'inférence pour laquelle on adhère à une hypothèse au détriment d'autres par son pouvoir explicatif des faits constatés. Certains auteurs (Kapitan, 1992) résument cette démarche comme celle d'une recherche de la meilleure explication.

Ce raisonnement est mis en place lors de l'utilisation de la formule de Bayes, en effet, admettons un ensemble d'hypothèses (H_1, …, H_n,) formant une partition dans l'espace de possibles modèles explicatifs d'une réalité, chacune de ces hypothèse est sensée expliquer un phénomène B, si B est plus probable de se manifester sous H_j que sur les autres H_k (k≠j), la réalisation de B renforce la conviction sur H_j. En termes symboliques :

$$P_B(H_j) = \frac{P_{Hj}(B)}{P(B)} \times P(H_j) \qquad (1)$$

dont :

$P(H_j)$: la probabilité portée sur l'hypothèse H_j avant l'arrivé du fait B; avec i = *1, …, j,…, n* (probabilité *a priori*)

$P_{Hj}(B)$: la probabilité de B sous l'hypothèse H_j

$P(B)$: la probabilité globale (ou totale) d'arrivée du fait B $(\sum_{i=1}^{n} P_{Hi}(B) \times P(H_j))$

$P_B(H_j)$: la probabilité portée sur l'hypothèse H_j après avoir appris B (probabilité *a posteriori*)

Pour une probabilité $P(B)$ donnée, le taux de changement entre $P(H_j)$ et $P_B(H_j)$ en (1) est donné par $P_{Hj}(B)$. Observons que plus probable est le fait constaté (*B*) sous une hypothèse *j* ($P_{Hj}(B)$), plus le degré de certitude augmente sur cette hypothèse :

Si $P_B(H_j) - P(H_j) > P_B(H_k) - P(H_k)$

C'est parce que

$P_{Hj}(B) > P_{Hk}(B)$

La méthode est un outil normatif de réassignation de mesures de certitude par le pouvoir explicatif des hypothèses. Remarquons que cette méthode ne permet pas la création spéculative abstraite des hypothèses ($H_1, H_2, \ldots, H_n$) ; processus qui semble complexe et soumis à la créativité humaine. En tout cas, cette méthode est une réponse objective (par delà son caractère normatif) à l'intention de tester des possibles théories ($H_1, H_2, \ldots, H_n$) par son accord avec les faits constatés (*B*).

Ce type de raisonnement est mobilisé dans un grand nombre d'activités humaines (argument fortement défendu par les bayésiens, certes), sa structure peut être trouvée dans certains modèles d'apprentissage en intelligence artificielle, dans la détection de courriels non désirés, de diagnostiques médicaux, etc.

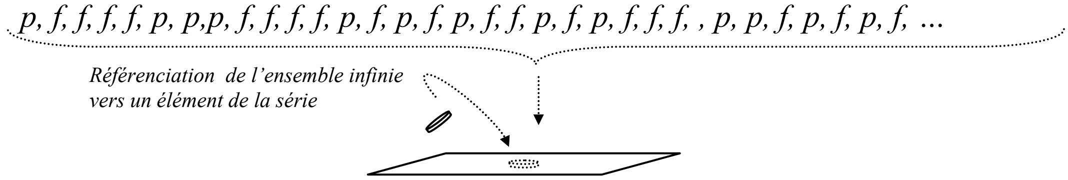
Or, celui-ci n'est pas le seul type de raisonnement mobilisé par la probabilité bayésienne. Il y'en a un autre, que nous avons appelé *référenciation* et qui apparaît lors des *épreuves génériques*.

La référenciation

Nous considérerons comme *référenciation* un type de raisonnement caractérisé par la recherche d'une classe de référence menant à la quantification de la probabilité. Ces classes nous les avons vues, sont de deux types; d'une part elle peut être celle de cas favorables dans les cas possibles, d'autre part, la fréquence d'apparition (Figure 4). Nous parlons de référenciation d'un élément par une classe pour décrire la mise en place d'un de ces deux principes permettant la quantification de la probabilité.

Deux types de principes dans le raisonnement par référenciation

Ensemble infini évoqué comme motif à croire



Ensemble fini évoqué comme motif à croire

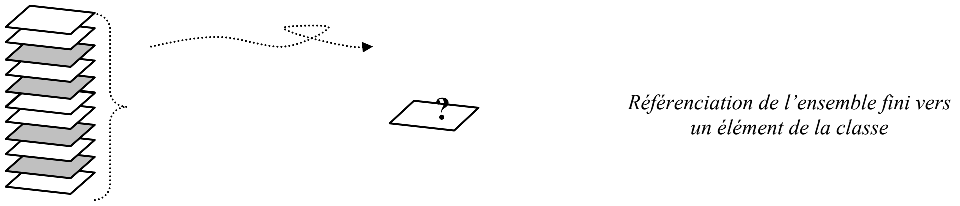


Figure 4

Les types de raisonnements par référenciation et par abduction restent cohérents avec les autres éléments caractéristiques de l'interprétation bayésienne :

- ils assignent des valeurs de vérité dans l'intervalle [0 ;1], qui s'expriment sous la forme d'une probabilité.
- Les propositions (ou objets) sur lesquelles se réalise cette assignation sont des *hypothèses* pour l'un et des *épreuves génériques* pour l'autre.

Remarquons que tous ces raisonnements (déduction, abduction et référenciation) ne sont pas mobilisés de manière isolée, l'abduction par exemple imbrique une déduction (ou référenciation selon le contexte) pour les évaluations des $P_{Hj}(B)$ et dans les cas les plus élémentaires une référenciation pour quantifier les probabilités *a priori* $P(H_j)$ dans une première itération de l'application de la formule de Bayes. Nous reviendrons plus tard sur ce sujet (*Chapitre I. Quatrième différence : Principes de quantification*).

En situations typiquement "scolaires" (lancers de dés, etc.) la référenciation apparaît comme naturelle ou même automatique; néanmoins dans le *Chapitre III* dédié à nos expérimentations, nous verrons que le recours à d'autres principes est source de conflits et d'opinions non conciliables chez les élèves et principalement lorsqu'une justification est sollicitée par l'enseignant; des conflits obligeant au professeur d'intervenir et d'institutionnaliser les principes de quantification en leur conférant un statut valide pour débloquer le conflit, principalement pour le principe de raison insuffisante.

Au présent, nous ne sommes pas en conditions de préciser dans quelle mesure ces principes d'évaluation doivent être explicités en classe. Néanmoins, nous verrons que d'un point de vue didactique, elles deviennent indispensables par leur fonctionnement comme variables didactiques favorisant l'émergence d'un ou autre type de raisonnement et son principe quantificateur respectif. Nous y reviendrons lors du troisième chapitre.

Nous venons d'esquisser les trois types de raisonnements mobilisés en fonction de l'objet sur lequel porte la probabilité et pour autant en fonction de l'interprétation de la probabilité en jeu. Ainsi, si l'on probabilise sur une hypothèse à partir des éléments factuels, nous sommes en présence de l'*abduction* (probabilité bayésienne) : sur une épreuve générique cela serait un raisonnement de type *référenciation* (probabilité bayésienne) et finalement une *déduction* lorsque l'objet est la série infinie (probabilité fréquentiste). Nous venons donc d'incorporer une troisième différence entre les interprétations de la probabilité. Le Tableau 2 le résume :

Trois des différences entre les interprétations de la probabilité

		Interprétations		
		Fréquentiste	Bayésienne	
Différences	Objet sur le quel probabiliser	Série infinie	Epreuve générique	Hypothèse
	Valeurs logiques portés sur ces objets	Vrai- faux	Intervalle continue [0,1]	
	Type de raisonnement	Déduction	Référenciaiton	Abduction

Tableau 2

Nous continuerons avec l'analyse des interprétations de la probabilité cette fois-ci dans un ouvrage de Laplace. Elle nous permettra aussi de positionner la définition de la probabilité classique en relation à la dualité d'interprétation de la probabilité.

1.8 Pierre-Simon Laplace. Une définition opératoire

Du vaste travail de Laplace nous nous sommes focalisé sur deux axes, le premier consiste à repérer le(s) sens donnée(s) par cet auteur à la probabilité, le deuxième à trouver une place à la formule classique de la probabilité dans l'approche de dualité de probabilité.

Pour la recherche des traces de la dualité de la probabilité chez Laplace nous nous sommes basés principalement sur le livre intitulé *Essai philosophique sur la probabilité* (Laplace, 1795).

Dans la première partie de cet ouvrage Laplace présente les généralités de son approche, dans la deuxième il en propose des applications telles que les jeux de hasard, les sciences naturelles et "morales" et même les jugement de tribunaux. Bien que la formule de cas favorables sur cas possibles est présente dès de début de l'ouvrage, nous avons commencé l'analyse des traces de la dualité chez cet auteur.

Le chapitre consacré aux sciences naturelles est assez proche de l'interprétation fréquentiste. Les exemples sont en général non seulement plausibles de reproduction ; Laplace évoque constamment la question du recueil d'un grand nombre d'observations pour étudier les lois de certains phénomènes ; il a même esquissé quelques éléments d'une théorie des erreurs de mesure. Nous sommes clairement dans une approche fréquentiste. Le contexte est celui de l'attente à long terme, même le théorème de Bernoulli est cité dans ce chapitre de son essai.

Pour ce qui concerne l'interprétation bayésienne, ses contextes typiques sont présents dès le début de l'ouvrage. La formule de la probabilité totale y est un outil récurrent pour Laplace. Les exemples d'urnes de compositions inconnues (probabiliser sur une hypothèse) sont fréquents et l'estimation de la probabilité d'un événement futur par un événement du passé (formule de probabilité totale) est récurrente. En fait Laplace s'est appuyé sur ces formulations mathématiques pour défendre sa conception déterministe de l'univers ((Laplace, 1795), page 3):

> *« (...)Nous devons donc, envisager l'état présent de l'univers, comme l'effet de son état antérieur et comme la cause de celui qui va suivre(...) ».*

La probabilité des "causes", autrement dit des hypothèses, est récursivement considérée pour modéliser des situations assez diverses. D'ailleurs, dans un passage intitulé « *Des inégalités inconnues qui peuvent exister entre les chances que l'on suppose égales* » (ibid. page 69) l'auteur s'intéresse aux conséquences que pourraient avoir les informations différentes sur les valeurs des probabilités. Nous ne prendrons qu'un seul exemple de la notion bayésienne chez Laplace, bien qu'ils soient nombreux tout au long de son essai.

Des témoignages

Le cas que nous avons retenu s'occupe de la probabilité de témoignages. Tous les ingrédients y sont présents pour une probabilité bayésienne. Ce genre de contexte s'avère propice pour cette notion, ci-contre un exemple d'un cas dont la probabilité modélise un rapport épistémique (ibid. page 124) :

> *«(...)la plupart de nos jugements étant fondés sur la probabilité des témoignages, il est bien important de la soumettre au calcul* (…) *on peut dans plusieurs cas, résoudre des problèmes qui ont beaucoup d'analogie avec les questions que l'on se pose, et dont les solutions peuvent être regardées comme des approximations propres à nous guider(...) ».*

Laplace propose ici le cas d'un "témoin" qui déclare avoir tiré[8] le numéro 79 d'une urne renfermant mille numéros. On s'intéresse à la probabilité de cette sortie, autrement dit, la probabilité que le numéro extrait soit effectivement le 79 déclaré par le témoin, celui-ci n'étant pas tout à fait crédible : *« l'expérience a fait connaître que ce témoigne trompe une fois sur dix ».*

La modélisation proposée par Laplace fait apparaître la véracité du témoigne comme un facteur pondérateur (9 /10) sur la probabilité d'avoir tiré le numéro 79 (1/1000) donnant finalement une probabilité de 9/10000 que le numéro tiré soit effectivement 79. Si le témoin trompe en nous disant que le 79 est sorti, elle sera de 999/1000 x 1/10.

En suite Laplace affirme que la crédibilité du témoin peut se voir engagée par des enjeux particuliers à la situation (*« avait quelque intérêt à choisir le numéro 79 »*) et ainsi se voir substantiellement modifiée (ibid., page 127) :

> *«(...) ½, 1/3, etc., suivant l'intérêt qu'il aura d'annoncer ce numéro. En la supposant 1/9, il faudra multiplier par cette fraction la probabilité 999/1000, pour avoir dans l'hypothèse du mensonge, la probabilité de l'événement observé (...) »*

[8] Evénement passé, la relation modélisée entre l'objet (numéro 79) et Laplace est du type épistémique

Nous voyons ici une autre caractéristique de la probabilité bayésienne qu'explique sa dénomination de "subjective", celle de varier en fonction des approches considérées. Pour Laplace, la probabilité peut, dans ce cas, prendre des valeurs différentes en fonction de l'intérêt du témoin pour annoncer le numéro 79.

Le caractère subjectif de cette probabilité est une conséquence du rapport épistémique qu'elle modélise, rapport qui devient plus évident à cerner si nous comparons la position du témoin avec celle de Laplace. Il peut arriver que ce témoin connaisse parfaitement le résultat du tirage en question. Néanmoins, le manque de connaissance de Laplace lui amène à traiter ce problème en termes de probabilités. La notion de rapport épistémique peut s'observer plus concrètement dans l'exemple suivant :

Deux élèves A et B sont assis tête à tête devant une table. On place une feuille de papier sur la table en mode de rideau (Figure 5). On lance une pièce de monnaie du côté de l'élève A et on leur demande à tous les deux de déterminer la probabilité que la face visible de la pièce soit pile. L'élève A ayant le résultat sur ses yeux répondrait 0 ou 1, tandis que l'élève B, se servant de quelques hypothèses habituelles en classe, évaluerait sa probabilité par un ensemble de référence, vraisemblablement sa probabilité serait de 1/2.

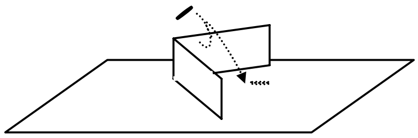


Figure 5

Peut-on considérer une probabilité fréquentiste ici ? Qu'elle soit éventuellement évoquée ne veut pas dire que cette interprétation soit la notion centrale. La fréquence participerait comme élément auxiliaire en tant que raison de croire, et en reprenant notre terminologie elle aurait le rôle d'*ensemble de référence infini* pour quantifier une *épreuve générique*, on serait ainsi en présence d'un type de raisonnement par *référenciation*.

En fait, il n'est pas clair que Laplace se serve de la rareté ou non des phénomènes évoqués comme raison, il cherche d'autres motifs, en général, ces exemples font appel au rapport entres les ensembles de cas favorables sur les cas possibles pour évaluer une probabilité, en d'autres termes un ensemble de référence fini ; démarche qu'il justifie comme étant naturelle (ibid., page 28) : *« Cette règle conduit à des résultats conformes aux indications du sens commun »*. Il est évident que ces exemples vont vers la direction de l'emploi de la formule. La plupart du temps, Laplace accompagne les exemples par un ensemble fini de référence, ce qui lui permet d'illustrer l'utilité de la formule. Ces cas seront en général des contextes du type épreuve générique référenciés par des ensembles finis.

Nous constatons donc encore une fois la présence de la dualité d'interprétation chez un même auteur, cette fois-ci chez Laplace et en particulier dans son ouvrage *Essai philosophique sur la probabilité*. Nous profitons de l'attention qu'il y a consacré à la formule menant à la quantification de la probabilité pour préciser nos raisons nous conduisant à ne pas la considérer comme une interprétation.

Quelle place pour la définition classique ?

Les interprétations (frequentiste et bayésienne) ont un statut bien différent de celui de la définition classique de la probabilité. Pour mieux comprendre cette différence, nous introduisons les concepts de *définition sémantique* et *définition opératoire*.

Définitions sémantique et opératoire

Dans le contexte de notre recherche, nous dirons qu'une définition est *sémantique* si elle rend compte du signifié d'un terme, ou en autres termes lorsqu'elle explicite l'empreinte signifiante du concept en question.

Une définition sera *opératoire* si elle détermine les étapes à suivre dans un procédé donné, dans ce sens une recette peut être considérée comme une définition opératoire, un algorithme en est un autre exemple.

Les *définitions* de la probabilité esquissées en la section 14 (fréquentiste et bayésienne) vont dans la direction d'une définition *sémantique*. Notons que dans ces deux définitions (même provisoires) se précise le signifié du terme probabilité :

«… La probabilité est alors la limite de la proportion de l'apparition de l'événement en question… » et *« …Elle exprime une mesure de certitude ou un degré de croyance ou de crédibilité portée sur une proposition…»*

La définition de la probabilité comme un quotient de cas favorables sur cas possibles est une définition *opératoire* parce qu'elle n'indique qu'un algorithme. Cette définition, dépourvue de sens et en tant que procédé de quantification, ne précise pas ce qu'on doit entendre par probabilité. Ce qu'elle nous indique sont des pas bien précis pour la quantifier. De plus, elle est utilisée par les deux paradigmes probabilistes, ce qui prouve son manque de signifié.

On peut aussi comparer ces deux type de définitions selon trois axes que nous appellerons : fonctionnel, hiérarchique et extensif.

- L'axe fonctionnel. Il nous parle des fonctions remplies par les définitions sémantiques et opératoires. Comme nous l'avons précisé ci-dessus, les définitions sémantiques ont l'intention de spécifier le signifié attribué à un terme. La définition opératoire a pour fonction de fournir une méthode de quantification.
- L'axe hiérarchique. Les unes habillent de signifié un mot, l'autre reste sous-ordonnée aux interprétations source et ne fait que préciser comment procéder pour passer au symbolisme mathématique. Le danger de mettre au premier plan la formule de Laplace est, parmi d'autres, celui de la perte de signifié des objets statistiques, et dans cette perte il y a au moins deux risques en statistique ; l'un consiste à faire croire que le travail statistique se résume à la manipulation mathématique d'objets dépourvus de signifié ; l'autre à en utiliser la symbolisation mathématique comme pivot pour échanger les interprétations de la probabilités (Figure 5) ((Jaynes, 1984);(Hacking & Dufour, 2004; Lecoutre, 2005)). En fin, nous parlons de hiérarchie parce que les définitions sémantiques se trouvent au dessus des définitions opératoires : une probabilité change de valeur par les opérations mathématiques effectuées, mais les traces de son signifié restent invariantes, tout au long de ces transformations.
- Extension. La dite définition opératoire, n'est pas valable pour tous les cas : ni pour la probabilité fréquentiste ni pour la probabilité bayésienne. On ne peut pas se contenter de la caractérisation de la formule de Laplace pour la substituer aux interprétations de la probabilité. Qu'en serait-il de toutes les situations fréquentistes où on ne peut pas la décomposer en cas élémentaires de même possibilité élémentaire ? Qu'en serait-t-il de toutes les situations bayésiennes dont les informations disponibles n'amènent pas à une distribution a priori équirepartie ? Et, comment considérer une probabilité que ne répond plus à ce schéma lors des transformations mathématiques ? Si la probabilité *est* un quotient entre cas favorables et possibles, alors la probabilité *est* un nombre[9] ((Keynes, 1921), page 22), ce qui est faux :

  « (...) It is often said, for instance, that probability is the ratio of the number of "favourable cases" to the total number of "cases" If this definition is accurate, it

[9] Nombreuses sont les situations où lorsque μ=p, l'auteur parle de μ comme étant une probabilité. Par exemple : la proportion de vote avec la probabilité de vote.

follows that every probability can be properly represented by a number and in fact is a number (...) ».

Erreur par échange d'interprétation postérieur aux calculs

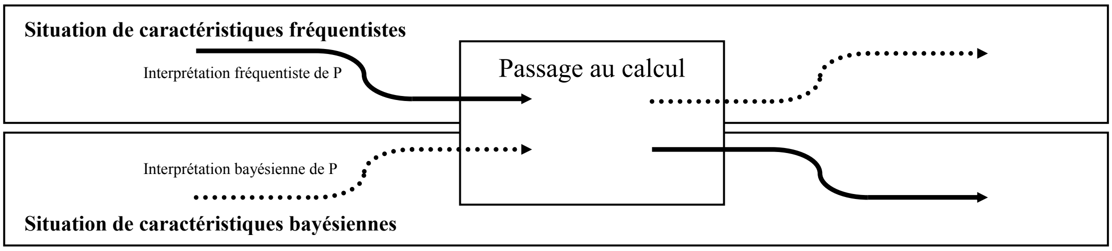


Figure 6

Bref, la place d'une définition *opératoire* de la probabilité en statistique en tant que quotient de cas favorables et cas possibles ne relève pas du même statut que les définitions *sémantiques* et dans le reste de notre travail, cette dernière ne sera pas prise en compte comme une interprétation en soi mais comme un passage (un parmi d'autres) à la modélisation mathématique. Après et en nous intéressant à une optique didactique, ces éléments seront repris du point de vue sémiotique.

Nous venons de placer la formule de Laplace ou définition classique de la probabilité comme algorithme permettant à la probabilité le passage au nombre. Dans ce contexte, elle n'est pas une interprétation dans le sens détaillé dans *Chapitre I. Définitions sémantique et opératoire.*

Nous comptons déjà avec trois types de différences entre les deux interprétations de la probabilité : les objets sur lesquels probabiliser, leurs valeurs logiques et le type de raisonnement mobilisé dans chacune des interprétations.

Dans la section *Chapitre I. Pour la probabilité bayésienne* nous avons parlé de la *référenciation* comme un procédé pour les *épreuves génériques* et aussi de la mise en place de deux principes permettant à la probabilité le passage au nombre, tous les deux utilisaient des ensembles de référence, l'un fini (formule de Laplace), l'autre infini.

Il nous manque d'autres principes qui facilitent le passage au nombre. Cela concerne toujours la probabilité bayésienne mais maintenant, c'est lorsque nous évaluons une probabilité dite *hypothèse*. Ce principe a lieu lors des premiers pas de l'*abduction* et il est couramment connu comme *principe de raison insuffisante*. Nous le développerons à partir des travaux de Keynes et ses réflexions sur la Logique inductive dans les années vingt.

Chez Keynes nous ne prétendons pas trouver des traces de la présence de la dualité, il fut un défenseur de l'approche bayésienne, principalement à partir d'une optique logique, nous retiendrons quelques-uns de ses apports à l'approche bayésienne de la probabilité.

1.9 John Keynes. Une approche logique

Keynes a beaucoup attiré l'attention dans les premières décennies du vingtième siècle par sa théorie de la probabilité comme une relation logique entre deux propositions. Un de ses livres les plus importants a été *A treatise on probability* (Keynes, 1921). Influencé par Leibniz, il est resté très attaché à la recherche de fondements logiques d'une probabilité caractérisée comme un degré de croyance rationnelle en la vérité d'une proposition (ibid., page 2) :

> *«(...) The subject matter of this book was first broached in the brain of Leibniz, who, in the dissertation written in his twenty-third year on the mode of electing the Kings of Poland, conceived of Probability as a branch of Logic(...)»*

Opposant radical de l'approche fréquentiste, il a conçu une sorte de logique inductive dont la probabilité est une mesure de crédibilité que l'on accorde à la validité des propositions dont l'intelligence renvoie à des connaissances limitées. Pour lui, les termes "certain" et "probable" décrivent des degrés de croyance rationnelle sur une proposition, ces degrés pouvant différer en fonction de la connaissance disponible. Bien que son idée soit proche de la probabilité conditionnelle, il s'est servi d'une autre nomenclature (ibid., page 4) : a/b = α, son corrélé en probabilité serait $P_b(a) = \alpha$.

Il a insisté sur l'existence d'une relation logique, objective et quantifiable, entre deux propositions là où il n'est pas possible de relier déductivement l'une à l'autre. Keynes et d'autres tels que Jeffreys ont insisté sur l'idée que les probabilités inductives renvoient aux éléments d'évidence, ils ont parlé de la relation entre la proposition *a* et l'évidence *b* plutôt que de la proposition *a* elle-même ; l'une (*a*) ne pouvant pas se considérer isolée de l'autre (*b*). Leur sujet d'intérêt n'était pas tant la probabilité de *a* mais la probabilité de *a* relative à *b*. Cette probabilité qui s'applique parmi d'autres à des hypothèses avait besoin de principes d'évaluation autres que celui d'équipossibilité, Keynes fut un des premiers à proposer celui de raison insuffisante.

Le principe de raison insuffisante

Comme tant d'autres, Keynes s'est vu confronté au problème de l'évaluation (ibid., page 20), mais pour lui cette valeur (α) est, nous insistons, une valeur de la relation logique entre deux propositions *a* et *b*. C'est ainsi qu'en s'intéressant à la probabilité portée sur des propositions il introduisit ce que aujourd'hui nous connaissons sous le nom de *principe de raison insuffisante* (ibid., page 42) :

> *«(…) The Principle of Indifference asserts that if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability. Thus equal probabilities must be assigned to each of several arguments, if there is an absence of positive ground for assigning unequal ones (…)»*

La justification de l'application de ce principe peut être très bien résumée par la phrase d'introduction du chapitre IV intitulé *The principle of indifference* (ibid., page 41) :

> *«(…) "ABSOLUTE. Sure, Sir, this is not very reasonable, to summon my affection for a lady I know nothing of" (…)*
> *"SIR ANTHONY. I am sure, Sir, 'tis more unreasonable in you to object to a lady you know nothing of" (…) ».*

Ce principe établit que devant un manque de connaissance sur la vérité d'un ensemble de propositions il n'y a aucune raison de privilégier l'une sur l'autre et pour autant, l'assignation de mesure de certitudes doit être uniforme, quelque chose comme une prise de position prudente face à l'inconnu.

Ce principe est de grande utilité aux bayésiens, il donne des raisons pour quantifier des probabilités là où on n'a aucune information significative sur les hypothèses en question, nous nous sommes centré sur ce principe dès un point de vue statistique.

Une généralisation de ce principe, même pour un ensemble discret d'hypothèses est celui de maximisation de l'entropie utilisée en théorie de l'information. A partir de cette approche, ce principe consiste à obtenir une distribution qui maximise l'entropie de l'information en fonction de l'information du système. Sans aucune information, la seule contrainte devient la somme de probabilités égale à un et la distribution qui en résulte est l'uniforme ((Jaynes, 1995) page 258).

Le principe de raison insuffisante vient en compléter d'autres que nous avons déjà présentés dans la section *Chapitre I. La référenciation* où nous avons décrit deux principes permettant de quantifier la probabilité des *épreuves génériques*. Nous intégrons ce nouveau principe aux autres déjà décrits et présentons une nouvelle différence basée sur les principes de quantification mobilisés pour chacune des interprétations de la probabilité.

Quatrième différence : Principes de quantification

Nous décrirons ce processus de quantification pour les deux interprétations de la probabilité pour ensuite revenir sur un tableau résumeant les différences entre les interprétations de la probabilité.

Probabilité fréquentiste

La probabilité fréquentiste est évaluée numériquement par deux grands moyens. L'un consiste en une estimation par la proportion d'apparition de l'événement en question sur un échantillon donné. L'autre, en admettant l'équipossibilité de cas élémentaires, il est possible de déterminer *a priori* la valeur de la fréquence d'apparition de l'événement.

Probabilité bayésienne. Épreuves génériques

Rappelons-le encore une fois, pour les *épreuves génériques,* nous comptions sur deux principes, tous les deux basés sur la référenciation par ensembles. Ces ensembles de référence peuvent être de deux types : l'un est l'ensemble fini constitué par les cas favorables dans des cas possibles, l'autre, celui de la série infinie, autrement dit la fréquence d'apparition.

Probabilité bayésienne. Les hypothèses

De notre parcours historique, nous avons retenu un principe pour évaluer les probabilités portant sur des *hypothèses,* en particulier lorsque l'on est en situation de manque de connaissance. Keynes[10] s'est occupé de ce sujet et a proposé un critère appelé *principe de raison insuffisante* qui donne des raisons pour distribuer uniformément les probabilités sur des *hypothèses*.

Si nous observons le procédé abductif de la formule de Bayes ne sert qu'à *réassigner* les probabilités sur les hypothèses plausibles $(H_1,...,H_n)$ lorsque l'on apprend une certaine information (B). Cette formule n'établit pas de critères pour les évaluer *a priori*, avant

[10] Feienberg ((Feienberg, 2006)) signale que le même principe à été proposé par Laplace en ((Laplace, 1771))

l'arrivée de l'information. Le principe de Keynes est un des critères les plus utilisés en probabilité bayésienne pour cette première étape du cycle de Bayes lorsque on ne dispose d'aucune information sur le système, nous présentons un exemple de son utilisation.

Un exemple d'inférence bayésienne

Hakima, l'enseignant, a sur sa table une bouteille opaque de bouchon transparent[11] et une urne remplie de billes noires et oranges. Hakima tire quatre billes de l'urne et les introduit dans la bouteille. Les élèves n'ont pas vu les couleurs des ces quatre billes, ils doivent estimer la composition de la bouteille. Hakima annonce que pour faciliter la tache il retournera la bouteille une fois et autre fois en laissant entrevoir la couleur d'une bille à travers le bouchon transparent mais il annonce qu'il demandera d'estimer la composition de la bouteille en termes de probabilité avant chaque retournement.

Une modélisation de ce problème par l'approche bayésienne mènerait à associer chaque composition possible à une hypothèse $(H_1,...,H_5)$ et les couleurs des billes observées à l'information disponible sur le système (B). Ainsi, chaque retournement de la bouteille renvoierait à une itération de la formule de Bayes (1), tant d'itérations comme retournements de la bouteille fasse Hakima.

$$P_O(H_i) = \frac{P_{Hi}(O)}{P(O)} \times P(H_i) \qquad \text{avec i=1 à 4} \qquad (1)$$

Dans ce problème, la couleur de la bille est apportée par Hakima, cette information serait utilisée pour réviser les degrés de certitudes portées sur les hypothèses, la formule de Bayes fourni des critères objectifs pour cette révision, néanmoins il faut encore assigner les probabilités sur les hypothèses $P(H_i)$ avant le premier retournement de la bouteille. Pour évaluer ces hypothèses il y a plusieurs possibilités, les voici quelques-unes :

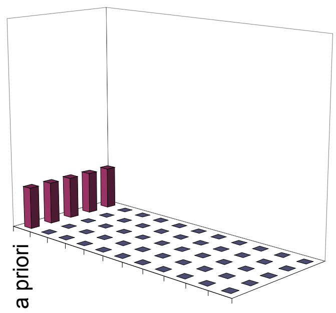
- Si la composition de l'urne est disponible, cette information permettrait aux élèves la représentation d'un contexte du type épreuve générique par la disponibilité d'un ensemble de référence (fini ou infini).

[11] Exemple inspiré d'un problème proposé par Brousseau ((Brousseau, Brousseau, & Warfield, 2001)) mais qui est ici traité à partir d' une approche bayésienne.

- Si la composition de l'urne n'est pas disponible, la représentation du type épreuve générique serait bloquée par l'impossibilité de représentation de la plausibilité de reproduction de l'épreuve. De cette manière les élèves, en se représentant des hypothèses devraient faire appel à d'autres principes que les ensembles de référence pour l'évaluer. Un principe serait celui de raison insuffisante traduisant une position quelque part prudente face au manque de connaissance.

Admettons que Hakima souhaite faire émerger le principe de raison insuffisante. Pour cela il inhiberait la représentation du type épreuve générique en refusant de donner des informations sur la composition de l'urne. Admettons aussi que les élèves choisissent le principe de raison insuffisante comme critère d'évaluation des probabilités, la distribution serait celle du Graphique 1.

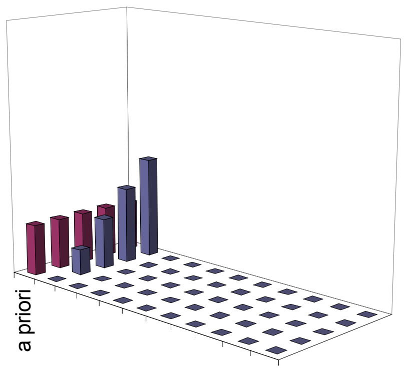
Distribution de probabilités *a priori* (principe de raison insuffisante)



Graphique 1

Un premier cycle peut se fermer lorsque Karima retourne la bouteille pour la première fois et obtient par exemple O (orange). Les probabilités a *posteriori* seraient alors calculées par la formule de Bayes. Le Graphique 2 représente l'évolution de la distribution de probabilités après avoir obtenu une bille orange dans le premier retournement, c'est la première itération.

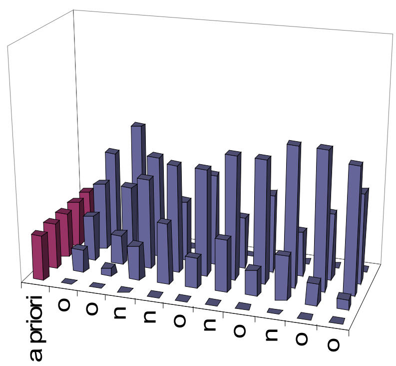
Distribution de probabilités *a priori* et *a posteriori* (première itération)



Graphique 2

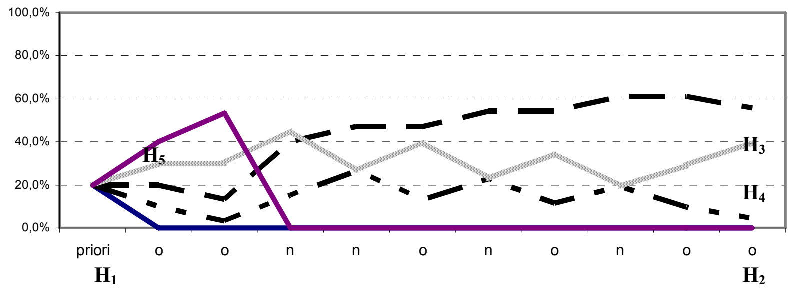
Admettons que Hakima réalise dix retournements de la bouteille et que les résultats sont : O, O, N, N, O, N, O, N, O, O. Les graphiques 3 et 4 montrent l'évolution des distributions de probabilité pour les compositions possibles (H_1 : NNNN ; H_2 : NNNO ; H_3 : NNOO ; H_4 : NOOO ; H_5 : OOOO).

Evolution de distributions de probabilités dix itérations (vue 1)



Graphique 3

Evolution de distributions de probabilités dix itérations (vue 2)



Graphique 4

Après ces dix retournements de la bouteille et en ne disposant d'aucune information sur système la composition la plus probable devient NNOO. Voici donc un exemple de transformation d'une probabilité sur le plan *opératoire* mais invariante sur le plan *sémantique*. Des probabilités qui portaient sur des hypothèses et évaluées par le *principe de raison insuffisante*. Et qui pour des raisons "complexes" et mathématiquement résumées par le théorème de Bayes subissent des transformations sur le plan *opératoire* mais qui n'ont jamais perdu les traces sur le plan sémantique. Une interprétation *sémantique* qui ne perd pas son empreinte même si elle change plusieurs fois de valeur dans une chaîne de dix itérations de la formule de Bayes.

En amont de cette lecture du problème en termes sémantiques et opératoires, nous avons voulu présenter ce test bayésien pour un ensemble discret d'hypothèses afin de montrer un contexte dont il est possible créer les conditions favorisant différents principes d'évaluation en fonction de l'information que nous mettons à disposition en classe. Pour cette occasion, le *principe de raison insuffisante* serait le visé, et cela à mode de variable didactique, en refusant toute information sur la composition de l'urne. Dans le *Chapitre III* nous analyserons un problème similaire, conçu pour inciter les élèves de BTS à l'utilisation de ce principe, nous verrons après en détail comment cette classe réagit à la nécessité de quantification des probabilités dans des contextes de manque de connaissance. Nous revenons sur notre recherche d'éléments caractéristiques tendant à différentier les deux notions de la probabilité.

Le contexte d'application de chaque principe

On pourrait supposer quelque part le *principe de raison insuffisante* contenu par la formule de Laplace, vu que les résultat en termes numériques sont les mêmes (une chance sur cinq à chaque composition). Néanmoins, les raisons menant à utiliser l'un ou l'autre ne sont pas du tout les mêmes, pour l'un c'est la connaissance d'information significative permettant d'accepter une hypothèses d'équiprobabilité ; pour l'autre c'est précisément le contraire, c'est le manque de connaissance qui amène à prendre une position initiale de prudence.

Les argumentations des élèves nous l'ont confirmé, chaque principe est avant tout une conséquence de la connaissance sur le système et non pas juste une valeur numérique, nous reviendrons sur ce sujet lors des analyses de nos expérimentations, nous continuons avec quelques remarques sur les objets pour la probabilité bayésienne.

Des liens entre les épreuves génériques et les hypothèses

Le cas décrit ci-dessus met en évidence la proximité des objets "épreuve générique" et hypothèse, ces objets typiquement bayésiens ont des similitudes et des différences, nous en précisons quelques-unes :

- les épreuves génériques et les hypothèses ne sont pas des compartiments clôt, leur délimitation parfois floue est en fonction des caractéristiques du problème.
- Une des principales différences concerne l'information fournie dans le contexte du problème.
- Un ensemble de référence dans le contexte facilite la plausibilité de reproduction de l'épreuve, cela mobilise soit le principe fréquentiste soit le principe d'équipossibilité.
- Si il n'y a aucun ensemble de référence, l'évaluation se voit forcée par d'autres critères, parmi eux, le principe de raison insuffisante.
- Cette délimitation en termes d'épreuves génériques et hypothèse a plutôt un intérêt didactique qu'épistémologique.
- Plausibilité de reproduction et principe mobilisé sont interdépendants : équipossibilité et fréquence d'apparition pour les épreuves génériques, raison insuffisante (parmi d'autres) pour les hypothèses.

Le type d'objet représenté (*épreuve générique* ou *hypothèse*) s'accompagne du principe mobilisé (ici *équipossibilité-fréquence* ou *raison insuffisante*), il n'y a pas l'un sans l'autre. Si nous pouvons nous représenter un objet générique, c'est parce qu'il y a un ensemble de référence qui facilite notre perception de la plausibilité de reproduction.

De même pour la représentation d'un objet type hypothèse, qui ne donnant pas l'image d'un objet dépourvu de singularité bloque les principes d'équipossibilité et fréquentiste et nécessitant d'une autre raison pour argumenter l'évaluation, ici le principe forgé pour Keynes.

Avec le *principe de raison insuffisante,* Keynes a cherché à objectiver la démarche d'assignation de probabilités et toujours à partir d'un point de vue logique. Son livre *A treatise on probability* est plein de réflexions très intéressantes : parmi elles, nous soulignons celles qui traitent sur des probabilités dont leurs comparaison est qualitative (ibid., *Chapitre III*) ou de ses critiques de l'approche fréquentiste.

Keynes n’a pas été le seul à s’intéresser à un versant logique d’une probabilité de racines bayésiennes : d’autres tels que Harold Jeffreys[12] ((Jeffreys, 1939) ; (Jeffreys & Wrinch, 1919)) contribuèrent à développer la probabilité bayésienne dans ses arêtes tant logiques que statistiques. Un autre exemple est Rudolf Carnap (Carnap, 1947) qui, avec son approche plutôt logique que statistique a développé un rigoureux système, toujours à partir de la notion de degré de certitude.

Quelle place pour la probabilité logique

On peut réunir ces quatre auteurs (Leibniz, Keynes, Jeffreys et Carnap) par leur intérêt commun envers les aspects logiques de la probabilité, bien que d’autres l’aient fait aussi (par exemple Ramsey en (Gärdenfors et al., 1988), page 19). Tous se sont intéressés à une probabilité en essence bayésienne, une probabilité interprétée d’une manière générale comme un degré de certitude, ou de croyance. Ils se sont intéressés également à la recherche d’une Logique fondée sur cette notion et avec leurs travaux, ils ont énormément contribué au développement de l’interprétation bayésienne de la probabilité en statistique comme l’atteste les concepts que nous avons retenus de Keynes.

Néanmoins, la probabilité dite "logique" entendue comme une notion fondamentale d’un champ scientifique spécifique tel que la Logique inductive ne nous semble pas devoir être considérée comme une interprétation de la probabilité en statistique. A partir de l’approche statistique qui nous intéresse, la probabilité logique sera pour nous considérée comme une des arêtes d’une interprétation de la probabilité, dans ce cas bayésienne. Nous poursuivons notre enquête épistémologique toujours centrée sur les interprétations de la probabilité qui sont arrivées à conformer un corpus de concepts leur permettant de bâtir un édifice inférentiel en statistique : les probabilités fréquentiste et bayésienne.

De notre autre objectif majeur pour ce chapitre, retenir des éléments caractéristiques aux deux interprétations de la probabilité, nous avons établit quatre différences, les différences abordées jusqu’à présent étant synthétisées dans le Tableau 3:

Quatre différences entre les interprétations de la probabilité

		Interprétations		
		Fréquentiste	Bayésienne	
fér enc	Objet sur le quel	Série infinie	Epreuve générique	Hypothèse

[12] En critiquant les idées Pearson ((Pearson, 1900))

	probabiliser					
	Valeurs logiques portés sur ces objets	Vrai- faux		Intervalle continue [0,1]		
	Type de raisonnement	Déduction		Référenciation		Abduction
	Quelques principes de quantification	Hypothèses du modèle	Equipossibilité (Ensemble fini)			Raison insuffisante

Tableau 3

Nous avons présenté très sommairement quelques idées de Keynes à partir de son livre intitulé *A treatise on probability* (Keynes, 1921). Dans cet ouvrage il a développé les concepts centraux de son projet consistant à établir les bases d'une Logique inductive ; la probabilité bayésienne étant une pièce fondamentale de cet édifice, elle est conçue comme objective. Néanmoins tous les statisticiens n'ont pas eu la même perception de la probabilité bayésienne. Pour de Finetti en particulier, la probabilité est aussi un degré de certitude mais incontestablement subjective. Nous analyserons quelques notions de son approche de la probabilité bayésienne et comme toujours, nous en tirerons profit pour continuer à caractériser les différences entre ces deux grandes écoles de la statistique inférentielle.

1.10 Bruno De Finetti. Une probabilité subjective

Les publications de de Finetti ont eu leur essor dans les années trente, nous nous centrerons sur quelques articles de cette époque et principalement sur une conférence qu'il a fait à l'Institut Henri Poincaré en mai de 1935 intitulée *La prévision : ses lois logiques, ses sources subjectives* (de Finetti, 1937). Dans son format papier, cette conférence est divisée en six chapitres, nous nous référerons aux deux premiers : *La logique du probable* et *L'évaluation d'une probabilité.*

Ses pensés sont explicitées dès le début même de la conférence (ibid., page 3) : « *Le point de vue que j'ai l'honneur d'exposer ici peut être considéré comme la solution extrême du côté du subjectivisme ».* Son subjectivisme consiste en ce qu'il n'y ait pas une graduation unique ou vraie pour la probabilité d'une proposition. Il remarquera que dans toute tentative d'évaluation de la probabilité, les choix sont nombreux et qu'il n'existe pas en soi une telle mesure théorique de la probabilité à partir de laquelle les autres ne seraient que des estimations. Pour de Finetti il y a juste une condition de suffisance à remplir pour les graduations de la probabilité, une condition logique, ou de cohérence en la quantification (ibid. page 8 et 16):

> *« (…) on entend que, une classe complète d'événements incompatibles E_1, E_2,… ; E_n étant donné, toutes les évaluations de probabilité qu'attribuent à p_1, p_2,…, p_n des valeurs quelconques non négatives et de somme égal à l'unité sont des évaluations admissibles : chacune de ces évaluations correspond à une opinion cohérente, à une opinion légitime en soi et chaque individu est libre d'adopter celle de ses opinions qu'il préfère, oui, pour mieux dire, celle qu'il sent (…)Y a-t-il donc parmi l'infinité d'évaluations parfaitement admissibles en elles-mêmes, une évaluation particulière qu'on puisse qualifier, dans un sens d'ailleurs encore inconnu d'objectivement juste ?Ou, du moins, peut-on se demander si une évaluation donnée est plus juste qu'une autre ? (…)»*

Cette conception de la probabilité bayésienne est parfois critiquée[13] : par exemple, si n'importe quelle distribution de probabilité est possible alors une prise de décision par le critère de maximisation de l'espérance peut conduire à une variété très large de décisions différentes en fonction des valeurs de probabilité choisies. Pour d'autres branches et dès un point de vue inférentiel, le caractère subjectif de la probabilité n'est pas un inconvénient majeur, ce qui en

[13] Ni la probabilité fréquentiste ni la probabilité bayésienne objective sont exemptes de critiques (Jaynes, Pearson, etc)

fait intéresse le plus c'est de disposer d'un critère objectif et normatif pour réassigner les probabilités lors de l'arrivée d'une information sur le système ((Gärdenfors et al., 1988), page 4) :

> *« (...) it's true that the "prior" probability values can be selected arbitrarily, but new information, obtained from experiments or by others means, will adjust the "subjective" probability distribution. One of de fundamental results in Bayesien theory, de Finetti's representation theorem, entails that even if two decision makers start out from widely different initial distributions (...) they will end up arbitrarily close to each other, if given sufficient times to experiment with it (...) »*

Pour montrer le caractère subjectif de la probabilité, de Finetti a proposé des exemples de procédés différents d'évaluation de la probabilité pour un même " événement". Un de ces exemples est celui de l'appréciation par la voie de classes différentes toutes contenant l'événement à probabiliser. Voyons que il s'agit d'évaluer la probabilité que : Monsieur A meure au cours de l'année, pour ce faire, on pourrait prendre la classe de personnes de même age, ou celle du même pays ou même encore celle des deux ensembles ou des classes plus restreintes définies par des individus en la même situation que Monsieur A (non fumeurs, végétariens, etc.). En appliquant le principe d'équipossibilité sur ces ensembles on peut s'attendre à des valeurs de probabilités différentes (ibid., page 21) :

> *«(...) Le choix d'une classe d'événements est en soi arbitraire, si l'on choisit des événements « semblables », ce n'est que pour rendre plus facile l'application du procédé (...) »*

Dans notre section *Chapitre I. Quelques précisions sur les épreuves génériques* nous avons déjà donné un exemple de calculs de probabilités différentes pour un même événement en fonction du choix de l'ensemble de référence. C'était le cas de Camille qui évaluait autrement que son père les probabilités de réussite à un examen. L'une le faisait en fonction de l'ensemble de jeunes candidats de banlieue, l'autre en fonction des candidats présentés à l'examen, tous confondus. Nous revenons sur l'article de de Finetti.

Pour rester cohérent avec sa subjectivité, de Finetti ne propose pas des critères spécifiques pour quantifier la probabilité. C'est ainsi que pour le passage au nombre, il donne une définition que contourne la question du critère de quantification (ibid., page 6) :

> *« (…) Supposons qu'un individu soit obligé d'évaluer le prix p pour lequel il serait disposé d'échanger la possession d'une somme quelconque S (positive ou négative) subordonnée à l'arrivée d'un événement donné, E, avec la possession de la somme pS ; nous dirons par définition que ce nombre p est la mesure du degré de probabilité attribué par l'individu considéré à l'événement E, ou, plus simplement, que p est la probabilité de E (selon l'individu considéré, cette précision pourra d'ailleurs être sous-entendue s'il n'y a pas d'ambiguïté)(…) »*

De Finetti affirme que tout critère qui respecte la condition de suffisance ($\sum_{i=1}^{n} P(H_i) = 1$) doit en principe être admis, et si il mentionne dans *La prévision* les procédés les plus courants, ce n'est que pour, en les comparant, mettre en évidence le caractère subjectif de la probabilité. Dans cette conférence en particulier, de Finetti s'est occupé de critères d'évaluation que nous avons appelé la référence par des ensembles finis et infinis ((de Finetti, 1937), page 22) :

> *« (…) ainsi le critère basé sur la notion de fréquence se réduit, comme celui des événements également probables, à une méthode pratique pour ramener certaines évaluations subjectives de probabilité à d'autres évaluations, elles mêmes subjectives (…) »*

Dans sa conception de la probabilité comme degré de certitude il a précisé que lorsque l'on parle de la probabilité de, par exemple, tirer "face", on le fait avec l'intention de s'intéresser à un coup en particulier (ibid., page 25) et que la considération des fréquences ou d'un ensemble qui contient l'événement n'est qu'un choix parmi d'autres visant une évaluation de la probabilité.

En fait, son concept d'épreuve nous semble proche de notre *épreuve générique* lorsqu'il précise (ibid., page 6) :

> *« (…) si l'on a à considérer plusieurs épreuves nous ne dirons jamais « épreuves d'un même événement » mas « épreuves d'un même phénomène » et chaque épreuve sera un événement (…) »*

Différencier entre une épreuve et la série que la contient est fondamental pour comprendre la notion d'épreuve générique. Nous avons introduit précisément cette dernière pour mettre en évidence que probabiliser par exemple sur le prochain tirage n'est pas une probabilité qui porte sur la série mais sur une " épreuve", une seule réalisation, même si à nos yeux elle ne se différencie en rien d'une autre de la même série, d'où la désignation de générique.

Des critères d'évaluation les plus courants, de Finetti en a fait allusion principalement à deux. L'un de ces principes, connu comme d'*équipossibilité,* donne des raisons à l'application de l'ensemble fini. L'autre principe donne des raisons pour se référencier à l'ensemble infini de la fréquence d'apparition de l'événement considéré. Il est connu comme *principe fréquentiste*. Nous consacrerons quelques lignes à ce principe permettant de quantifier la probabilité à partir de la fréquence d'apparition de phénomène considéré.

Le principe fréquentiste

Le *principe fréquentiste* est un critère d'évaluation pour des épreuves génériques et à la différence de celui d'équipossibilité, il le fait par la fréquence d'apparition de l'événement considéré. Ian Hacking le décrit comme ((Hacking & Dufour, 2004), page 149) :

> *« (...) Une espèce de règle (...) établissant un lien entre la probabilité fréquentiste et probabilité épistémique. Il s'agit d'une règle sur la connaissance et l'ignorance. (...)Si S est un événement particulier de type E. Si la seule information disponible sur E à l'issue d'essais d'un certain type, et réalisés selon une certaine procédure aléatoire, est que la probabilité fréquentiste de E est P(E)= p alors la probabilité épistémique de S vaut également p (...) ».*

En d'autres termes, ce principe établit que face à une situation de manque de connaissance sur une épreuve en particulier (au sens large) dont on ne connaît que la fréquence de son apparition, alors sa probabilité peut être évaluée par celle-là, et ainsi, plus fréquemment se produit l'évènement, plus on y croît ((Gärdenfors et al., 1988), page 40 et 110).

Le *principe fréquentiste* se confond avec celui de l'*équipossibilité* si les épreuves élémentaires sont équiprobables. Un grand nombre d'exercices scolaires appartient à cette catégorie où on peut évaluer la probabilité par le *principe d'équipossibilité* ou par *le principe fréquentiste*. Les exemples à partir d'urnes et de jeux de cartes en sont des exemples, tandis que l'exemple de la punaise (Quelques précisions sur les épreuves génériques) ne permet que la référenciation par le *principe fréquentiste* (Figure 7).

Quelques procédés d'évaluation de probabilité bayésienne pour épreuves génériques

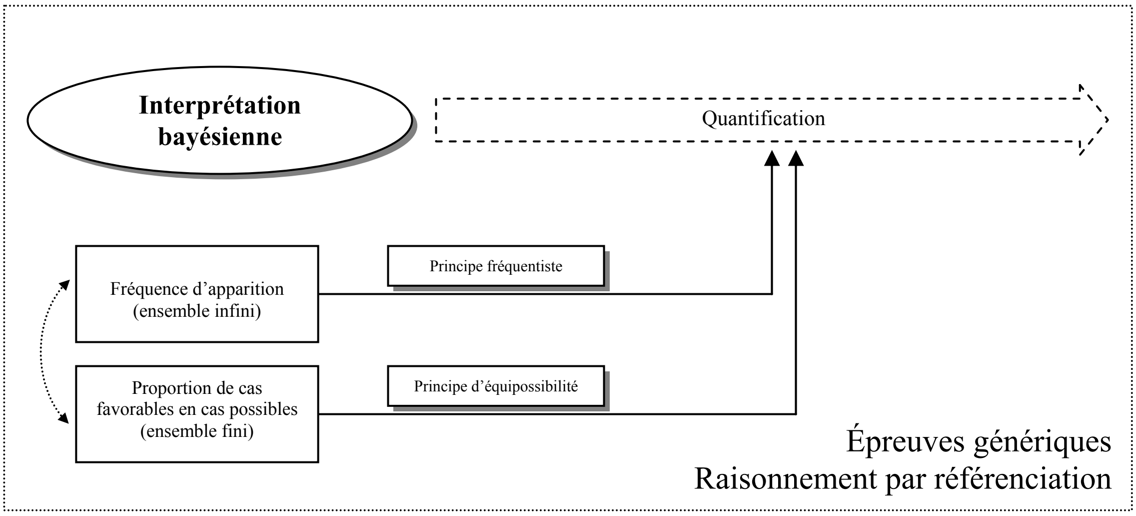


Figure 7

Le *principe fréquentiste,* permet de différencier le rôle de la probabilité fréquentiste lors de l'évaluation d'une épreuve générique. Une épreuve générique par définition ne possède pas de particularités et lorsque sa reproduction est possible, il est habituel de prendre la fréquence d'apparition (si possible) comme critère d'évaluation de la probabilité. La présence de cette fréquence peut amener à penser que l'interprétation en jeu est la fréquentiste mais en réalité, elle intervient dans un plan secondaire comme un principe de quantification ou d'évaluation. Le cas de la punaise où on s'intéresse à la probabilité d'une épreuve (le prochain tirage), est un exemple de la contribution de la fréquence à la quantification de la probabilité d'une épreuve générique.

Cette articulation entre la probabilité bayésienne et la fréquentiste par le principe fréquentiste en épreuves génériques sera reprise dans une deuxième étape (*Chapitre II*) pour caractériser les interprétations sous-jacentes dans les exercices proposés aux élèves dans les manuels de lycées français. Dans ce chapitre, nous cherchons à confirmer l'hypothèse de la présence incontournable de la dualité de la probabilité en amont des intentions des programmes et documents officiels de la présenter comme fréquentiste. Nous analyserons en plus les contextes des exercices avec les éléments caractéristiques identifiés dans ce premier chapitre pour dessiner des espaces potentiels de travail dans le chapitre Probabilité des manuels de lycée.

Avec le principe fréquentiste nous complétons la quatrième ligne de notre tableau correspondant aux différences entre les interprétations de la probabilité. Ce principe non seulement en est un de différenciation mais aussi un d'articulation entre les deux interprétations.

Quatre différences entre les interprétations de la probabilité

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="5">Interprétations</td></tr>
<tr><td colspan="2">Fréquentiste</td><td colspan="3">Bayésienne</td></tr>
<tr><td rowspan="4">Différences</td><td>Objet sur le quel probabiliser</td><td colspan="2">Série infinie</td><td colspan="2">Epreuve générique</td><td>Hypothèse</td></tr>
<tr><td>Valeurs logiques portés sur ces objets</td><td colspan="2">Vrai- faux</td><td colspan="3">Intervalle continue [0,1]</td></tr>
<tr><td>Type de raisonnement</td><td colspan="2">Déduction</td><td colspan="2">Référenciation</td><td>Abduction</td></tr>
<tr><td>Quelques principes de quantification</td><td>Hypothèses du modèle</td><td colspan="2">Equipossibilité (Ensemble fini)</td><td>Fréquentiste (ensemble infini)</td><td>Raison insuffisante</td></tr>
</table>

Tableau 4

Nous venons d'esquisser quelques idées centrales de deux écoles de la probabilité bayésienne, l'une caractérisée comme logique et objective (Keynes, Jeffreys, etc.) et l'autre comme subjective (de Finetti, Savage, Ramsey, etc.). En opposition à l'approche bayésienne, nous présenterons les concepts Richard von Mises, et de son frère Ludwin, tous les deux déclarés fréquentistes. Le premier a développé une théorie de la probabilité, elle nous semble, beaucoup plus solide que celle de Ludwin qui, en adhérant aux principes de celle de son frère Richard, s'est intéressé aux applications de la probabilité en économie. La théorie fréquentiste de Richard von Mises nous permettra de traiter la probabilité propensioniste de Popper et Peirce.

1.11 Richard von Mises. La probabilité fréquentiste

Une notion théorique

Richard von Mises a publié entre autres l'ouvrage *Probability, statistics and truth* (von Mises, 1928). L'auteur y développe sa théorie fréquentiste de la probabilité à partir de deux axiomes, le premier traite sur la convergence de la série infinie; le deuxième sur la randomisation des éléments de la série.

La position de von Mises est complètement différente de celle des bayésiens, il n'a pas tenté de modéliser le terme "probabilité" dans son usage courant ni non plus de comparer ses arguments avec ceux de "l'homme de la rue" comme le font quelques les bayésiens. Dès les premières pages, von Mises insiste sur le caractère scientifique de sa théorie en différenciant son signifié de la probabilité de celui de la vie courante. Il souligne d'ailleurs son non intérêt envers les problèmes des "sciences morales" tel que l'a fait Laplace ((von Mises, 1928), page 8) ; ceci afin d'éviter la probabilité appliquée sur des hypothèses.

La probabilité, affirme von Mises, est une notion scientifique et diffère de son homonyme de la langue comme le fait le concept de travail en Mécanique. Cette acception scientifique de la probabilité est définie par l'auteur comme la limite de fréquence d'apparition d'un événement donné. Afin d'éviter des ambiguïtés, von Mises délimite le type des situations auxquelles cette notion peut s'appliquer. Pour cela, il introduit un concept qui précède celui de la probabilité, il précise que la probabilité est définie sur un *"collective"*, sans cet ensemble, il n'y a pas de probabilité. Un *collective* (ibid., page 11) :

> *« (...) denotes a sequence of uniform events or processes which differ by certain observable attributes, say colours, numbers, or anything else (...) »*

Un *collective* serait une séquence d'événements ou occurrences capables, en principe, d'être continuées indéfiniment comme par exemple une séquence de lancers faite avec un dé supposé indestructible ((Popper, 1959), page 139) :

> *« (...) each of these events has a certain character or property; for example, the throw may show a five and so have the property five. If we take all those throws having the property five which have appeared up to a certain element of the sequence, and divide their number by the total number of throws up to that*

element (i.e. its ordinal number in the sequence) then we obtain the relative frequency of fives up to that element (...)»

En plus de la notion de *collective*, von Mises introduit deux axiomes dans sa théorie de la probabilité fréquentiste.

Les axiomes

Un *collective* doit satisfaire deux conditions formulées sous la forme d'axiomes :

- Axiome de convergence : La fréquence relative du caractère (ou propriété) observé dans le *collective* tends vers l'infini.
- Axiome de randomisation : La limite de convergence ne se voit pas affectée par la sélection d'un sous ensemble infini quelconque à la condition que la règle de sélection soit fixe. Si nous choisissons avec un critère fixe un sous ensemble infini de la série originairement convergente, alors la fréquence d'apparition du caractère en question convergera dans ce nouveau sous-ensemble vers la même valeur que dans la série d'origine.

Il y a donc deux conditions de nécessité (ou axiomes) pour la probabilité, si ceux-ci sont satisfaits, on serait devant un *collective.* Seulement, sous ces conditions, von Mises désigne par probabilité la limite de la fréquence d'apparition du caractère observé. En d'autres termes, l'objet sur lequel on peut probabiliser n'est ni plus ni moins que la *série à long terme*.

Von Mises a insisté sur le caractère objectif et unique de la probabilité pour un *collective* donné, une probabilité disposant de propriétés empiriques et conçue comme une magnitude à laquelle on peut se rapprocher de plus en plus par des expériences successives. Selon lui, il existe un tel concept de probabilité, et d'une manière analogue à la masse ou à une résistance électrique, la probabilité serait une propriété physique d'un *collective* dont les fréquences empiriques ne seraient que des approximations à cette propriété de la classe appelée *collective*.

Une probabilité pour la série infinie

Von Mises remarque que cette définition de la probabilité n'est pas applicable à ce que nous avons dénommé des *épreuves génériques*, même s'il existe un *collective* dans lesquels les inscrire. Par exemple, sur la probabilité de décès de Monsieur X il dira ((von Mises, 1928), pages 17-18) :

> « *(...) the 'probability' of Mr. X dying in the course of the next year, for instance (...) is "utter nonsense (...)* ».

Pour von Mises, ce n'est pas un problème de manque d'ensemble de référence ou de *collective,* bien au contraire en est un de cohérence. Pour Monsieur X il y a beaucoup de *collectives*, du même que pour Camille lorsqu'elle et son père cherchaient à évaluer les chances de réussite de son prochain examen. Pour le paradigme fréquentiste, probabiliser sur une épreuve générique renvoie à une contradiction, s'il y a plusieurs ensembles de référence ou *collectives*, chacun renvoierait à une mesure de probabilité différente et ainsi on serait en contradiction avec le caractère unique de la probabilité[14].

D'après Richard von Mises donc, ni une *épreuve générique* ni une *hypothèse* ne peuvent être soumises au calcul de probabilité (ibid., pages 12 et 33 respectivement) :

> *« (...) It is possible to speak about probabilities only in reference to a properly defined collective (...) The definition of probability (...) is only concerned with the probability of encountering a certain attribute in a given collective (...)»*

Ainsi le modèle proposé par Richard von Mises ne serait valable que pour des situations où la série infinie est explicitement l'objet à probabiliser et non pas sur des épreuves génériques où la série intervient comme un ensemble de référence. Pour lui, l'ensemble qui converge (*collective*) est le seul objet sur lequel on peut probabiliser (ibid., page 15) :

> *« (...) the relative frequency of the observed attribute would tend a fixed limit if the observations were indefinitely continued. This limit will be called the probability of the attribute considered within the given collective. This expression being a little cumbersome, it is obviously not necessary to repeat it always. Occasionally, we may speak simply of the probability of 'heads'. The important thing to remember is that this is only an abbreviation, and that we should know exactly the kind of collective to which we are referring (...)»*

Et à la question de la probabilité de décès d'un individu X en particulier, il répond (ibid, page 17) que seulement à partir d'une perspective générale (par exemple d'une compagnie d'assurance) et disposant d'une masse considérable de données, on peut estimer la fréquence de décès d'un ensemble d'individus, mais dans aucun cas sur Monsieur X en particulier.

[14] En fait, il y des raisons beaucoup plus profondes que cette cohérence, elles sont d'ordre philosophique concernant des sujets tels que la vérité, l'induction, et la découverte scientifique que nous ne traiterons pas dans notre synthèse.

Cette approche diffère profondément de celle des bayésiens en particulier les subjectivistes pour lesquels une probabilité unique et objective n'a pas de sens. Par exemple, de Finetti réplique que *« la probabilité n'existe pas »* ((De Finetti, 1974), page 5) en s'adressant à ceux qui défendent une approche de la probabilité comme étant unique et objective. De sa part, von Mises rejettera catégoriquement l'existence d'une probabilité *a priori* ainsi que ses respectifs principes d'évaluation (raison insuffisante, fréquentiste, équipossibilité, etc.). Une telle chose *a priori*, détachée de sa fréquence d'apparition n'existe pas pour Richard von Mises (ibid., page 71) :

> *« (…) How is it possible to be sure (…) that each of the six sides of a die is equally likely to appear? (…) Our answer is of course that we do not actually know this unless the dice…have been the subject of sufficiently long series of experiments to demonstrate this fact (…)»*

Quelques années plus tard, son frère Ludwin, en *Human Action : a treaitise on economics* (von Mises, 1966) propose autre définition de probabilité fréquentiste, à notre avis beaucoup moins précise.

Class probability et *case probability de Ludwin von Mises*

Cette quatrième édition de, probablement, l'oeuvre la plus connue de Ludwin von Mises comporte trente neufs chapitres. Le chapitre six de la première partie intitulé *Uncertainty* est consacré à la probabilité et plus particulièrement à quelques réflexions autour de « *human actions* » en contexte d'incertitude.

Si pour Richard von Mises la définition de probabilité concerne un ensemble infini satisfaisant les axiomes de convergence et de randomisation, pour son frère Ludwin les concepts ne sont pas si clairs. En fait Ludwin partage l'opinion de son frère sur la probabilité fréquentiste mais lui donne une approche personnelle afin de traiter le sujet qui vraiment les intéressent : une théorie sur l'économie. De toute manière, même si personnelle, son interprétation nous permet de mettre plus en évidence encore le champ d'application de la probabilité fréquentiste, en nos propres termes, l'objet sur lequel probabiliser.

Ludwin von Mises défini deux probabilités ((von Mises, 1966) page 106) :

« (...) There are two entirely different instances of probability; we may call them class probability (or frequency probability) and case probability (or the specific understanding of the sciences of human action)(...)»

Sur sa *class probability* il la décrit, tel quelle que son frère, à savoir comme une caractéristique d'un ensemble dont on connaîtrait tout sur le comportement d'une classe d'événements mais que sur un événement singulier on ne saurait rien sauf qu'il appartient à la classe. Ensuite des exemples plus où moins précis sont présentés :

- Il y a quatre vingt dix billets de loterie où cinq seront tirés. On admettrait tout connaître sur le comportent de l'ensemble des billets mais par rapport à un billet en particulier on ne connaîtrait rien sauf qu'il appartient à l'ensemble des billets.
- On dispose d'une table de mortalité définie sur une période de temps passée et sur une région déterminée. Si l'on assume qu'aucun changement n'arrivera, on peut dire que l'on connaît tout sur la mortalité de l'ensemble en question mais par rapport à l'attente de vie d'un individu on ne connaît rien sauf qu'il appartient à cette population.

A notre avis il peut y avoir ici une confusion entre la classe des événements à laquelle se référait Richard von Mises, (par exemple l'ensemble infini d'extractions avec remise de cinq billets de loterie) et la classe d'éléments à laquelle semble faire allusion Ludwin von Mises (l'ensemble fini de billets de loterie). En tout cas, tous les deux seront d'accord en ce que la probabilité fréquentiste ne porte pas sur un de ces éléments. A cet effet Ludwin de même que son frère affirme ((von Mises, 1966), page 108) :

« (...) [le calcul de probabilités] do not lead to results that would tell us anything about the actual singular events (...) The fruit dealer may know, for instance, that one of every fifty apples will rot in this stock; but he does not know to which individual apple this will happen (...) There are, of course, many instances in which men try to forecast particular future event on the basis of their knowledge about the behavior of the class. A doctor may determine the chances for the full recovery of his patient if he knows that 70 per cent of those afflicted with the same disease recover. If he expresses his judgment correctly, he will not say more than that the probability of recovery is 0.7, that is, that out of ten patients not more than three on the average die (...) ».

La différence (et qui ne sera qu'apparente) entre Ludwin et Richard von Mises réside en ce que le premier reconnaît une "probabilité" pour les événements singuliers mais, cette notion n'est pas scientifique ni soumise aux calculs de probabilité.

Enfin, dans ce qui peut paraître comme une différence conceptuelle entre les frères von Mises, il nous semble que ce n'est que de terminologie et que tous les deux ont partagé l'idée d'une probabilité représentant strictement la stabilisation de fréquences.

Les épreuves génériques ont posé de nombreux problèmes aux fréquentistes : très courantes en sciences, elles font partie d'une série qui répond aux conditions d'application de la probabilité fréquentiste mais son caractère unitaire empêche toute analyse en termes fréquentiste.

Karl Popper, fervent fréquentiste dans *The logic of scientific discovery* ((Popper, 1959)), a essayé de combler ce vide de la théorie fréquentiste en proposant une nouvelle interprétation de la probabilité, une sorte de variation de la probabilité fréquentiste appelée propensioniste, variation qui à mode d'ajustement permet, affirme Popper, de traiter les épreuves génériques, sans remettre en question son paradigme sur la découverte scientifique et la falsification.

1.12 Karl Popper. La probabilité propensioniste

Avec l'interprétation propensioniste, Popper a cherché dans *The propensity interpretation of probability* ((Popper, 1959)), à créer un cadre conceptuel pour traiter ce que nous avons dénommé les *épreuves génériques*. En effet, un arrangement dans l'interprétation fréquentiste semblerait lui suffire pour probabiliser sur ces épreuves plausibles de reproduction ; pour les autres, les *hypothèses* Popper sera intransigeant (ibid., page 26) :

> *« (…) There may be something like a measurable degree of the rationality of a belief in a, given the information b; but I assert that this belief cannot be adequately measured by a measure that satisfies the laws of the calculus of probability (…) »*

Popper souligne à plusieurs reprises son approche de l'interprétation de la probabilité (ibid., page 26) :

> *« (…) asserting nothing but that the relative frequency of the event [A] in a sequence defined by the conditions [B] is equal to [R]. Or in other words, the statement '$P_B(A) = R$' is interpreted to mean: " events of the kind [A] occur, in sequences characterised by [B], with the frequency [R]". Thus, for example, 'PB(A) = R' may mean "the relative frequency of tossing heads with a normal penny equals ½" (where [A] is getting heads upmost, and [B] is a sequence of tosses with a normal penny)(…) »*

Le paradigme fréquentiste n'admet autre chose que le *collective* de Richard von Mises. Pour probabiliser sur une épreuve générique, Popper tente de ranger la définition fréquentiste (ibid. pages 26 et 30 respectivement) :

> *« (…) I changed my mind in favour of the propensity interpretation (…) for two reasons:*
>
> *(1) The first was connected with the problem of the interpretation of quantum theory.*
>
> *(2) The second was that I found certain flaws in my own treatment of the probability of "single events" (in contrast to sequences of events), or "singular events" as I shall call them in analogy to "singular statements" (…) [la probabilité fréquentiste] do not allow us to predict, or to say, anything*

whatever about a single event, except that its repetition, under the same conditions, will generate a sequence with certain statistical properties(...) »

Par *« single events » Popper* entend (ibid., page 29) un événement unitaire tel que par exemple *« obtenir un six au troisième lancer fait à neuf heures du matin d'aujourd'hui avec ce dé »*. Ces événements, nous les avons dénommés comme des *épreuves génériques*. Popper propose donc la notion de *propension* pour traiter ce genre de situations. Avec ce concept, il tente de déplacer la propriété, qui pour les fréquentistes se trouve dans la série infinie, vers les conditions de l'expérimentation. De cette manière, on pourrait rescinder la série tout en gardant sa potentielle reproduction. Pour Popper, la propension serait une tendance, une disposition ou même une facilité d'un artefact à fournir tel ou tel résultat ; facilité ou disposition déterminée par les conditions de (re)production de l'événement.

La principale différence entre la notion fréquentiste et la propensioniste réside dans la localisation de la propriété, pour la première c'est dans la série, pour la deuxième c'est dans les conditions de production. De cette manière, et comme les conditions sont invariantes tout au long de la réalisation de la série d'épreuves, elles sont présentes aussi pour une de ces épreuves en particulier. Ainsi Popper déclare pouvoir probabiliser donc sur une telle épreuve ; la probabilité serait alors une mesure sur les conditions de production et non pas sur la série.

Pour cela, Popper a développé toute une élaboration métaphorique, en essayant de cerner le concept et en le différenciant au même temps de l'interprétation fréquentiste (ibid., page 30) :

« (...) the concept of propensity, or of a field of propensities, introduces a dispositional property of singular physical experimental arrangements—that is to say, of singular physical events—in order to explain observable frequencies in sequences of repetitions of these events. In both cases the introduction of the new idea can be justified only by an appeal to its usefulness for physical theory (...) But part of the usefulness of these concepts lies precisely in the fact that they suggest that the theory is concerned with the properties of an unobservable physical reality and that it is only some of the more superficial effects of this reality which we can observe, and which thus make it possible for us to test the theory (...)»

Sauf le déplacement de propriété de la série vers la disposition de l'artefact qui la produit, Popper garde toutes les caractéristiques de la probabilité fréquentiste et même sa valeur numérique. En effet, pour évaluer la probabilité propensioniste appliquée sur une épreuve générique, Popper tout simplement prend la propriété de la série convergente (ibid., page 33) :

« (…) we have no doubt that their singular probability is (…) estimated as being equal to the frequency of the sequence (…) »

Par exemple, si en nous servant d'un dé équilibré on s'intéresse à la probabilité d'obtenir un quatre lors du prochain lancer, l'évaluation serait, pour l'approche propensioniste, de 1/6. Cette valeur représenterait la disposition ou facilité du dé à faire sortir la valeur quatre.

Cette seule adaptation permet, d'auprès Popper, de combler le vide laissé par la probabilité fréquentiste. Le reste, et en particulier les procédés inférentiels resteraient sans modification pour cet auteur.

L'effort de Popper pour combler le vide de la notion fréquentiste tout en gardant son cadre général doit s'inscrire dans la défense d'un projet plus ambitieux, celui de postuler une théorie sur la découverte et la falsification d'une théorie scientifique (Popper, 1959). Par exemple, une probabilité propensioniste serait falsifiable par la répétition d'épreuves, on pourrait d'après Popper, falsifier une valeur R donnée d'une probabilité en accumulant des preuves contre la fréquence R.

Cette théorie falsificasioniste est aujourd'hui parfois contestée, elle semblerait ne pas toujours représenter la démarche de la découverte scientifique, quelques épistémologues (Hacking, 1965) insistent sur le fait que les scientifiques d'aujourd'hui cherchent non pas des preuves afin de rejeter une théorie (test fréquentiste ?) mais plutôt afin la confirmer (test bayésien ?).

Popper n'as pas été le seul à défendre une approche propensioniste. Plus ou moins proche, nous trouvons par exemple Charles S. Peirce qui lui aussi a suivi les pas de Popper en se déclarant dans un principe fréquentiste pour ensuite devenir propensioniste. De toute manière, cette tentative ne semble avoir été retenue par la communauté scientifique et celle-ci constitue la principale raison pour laquelle nous écartons cette approche de nos analyses des interprétations de la probabilité en statistique.

Même si nous ne la retenons pas, cette interprétation nous a permis d'enlever des doutes sur les caractéristiques des objets sur lesquels une probabilité fréquentiste peut porter. En montrant comment Popper s'est vu obligé de reformuler la notion de probabilité fréquentiste est que nous souhaitons préciser la nature des objets sur lesquels de la notions fréquentiste de la probabilité porte.

Ces travaux de Popper datent des années cinquante, Kolmogorov pour sa part s'est déclaré fréquentiste (Neyman, 1977), d'autres tels que Savage ou Jeffreys, bayésiens, etc. La période des années trente jusqu'aux soixante-dix a été de grandes controverses, les deux écoles se sont consolidées tant dans leurs aspects philosophiques que méthodologiques. Les développements informatiques des années quatre-vingt ont simplifié la charge mathématique des tests, néanmoins, il reste toujours sans automatiser le control sémantique des opérations réalisées. Aujourd'hui nous trouvons des paquets informatiques fournissant les résultats mathématiques d'un test dont son interprétation reste toujours sous la responsabilité de l'usager.

Notre intérêt s'adresse vers cette direction, l'interprétation des résultats d'un procédé. Des interprétations indispensables à l'argumentation d'une prise de décision. Dans ce sens nous insistons sur le fait que les tâches calculatoires sans l'habillage interprétatif ne constituent pas une argumentation et que sans ce dernier, toute modélisation d'une réalité devient inconsistante, perdant ainsi sa représentativité.

En ce qui concerne la statistique, la probabilité joue un rôle primordial dans une démarche de modélisation, ses interprétations déterminant des méthodes inférentielles bien différenciées. Dans ce premier chapitre nous avons montré plusieurs approches de la probabilité et en particulier celles des deux grandes écoles inférentielles connues comme fréquentiste et bayésienne. Nous avons aussi retenu des éléments caractéristiques à ces deux interprétations. Par la suite nous présentons un résumé de ce premier chapitre.

1.13 Conclusions

Nous avions deux objectifs principaux pour ce chapitre :

- Le premier, repérer les traces d'une dualité d'interprétation de la probabilité au long de son histoire, depuis son émergence au dix-septième siècle jusqu'à nos jours.
- Le deuxième, les deux interprétations de la probabilité, afin de compter sur des outils précis nous permettant de les identifier.

A ces deux objectifs correspondaient deux motivations didactiques :

- Pour le premier, interpeller la transposition didactique d'un objet de l'enseignement par une enquête épistémologique. Un objet présenté comme objet à enseigner fréquentiste
- Pour le deuxième, le repérage de différences, notre intention était de nous outiller le plus objectivement possible afin d'identifier la présence de la dualité dans les exercices de manuels (et ainsi confirmer notre hypothèse d'*incontournable* dualité) et en plus de compter sur des éléments pour élaborer nos expérimentations en classe.

Du repérage de la présence de la dualité au long de son histoire

Des incertitudes différentes

Nous avons commencé par Blaise Pascal et montré dans deux de ses problèmes la présence d'éléments distinctifs à chacune des interprétations de la probabilité. Ni les termes ni les symboles n'y étaient, néanmoins les situations abordées commençaient à dessiner une probabilité à deux visages. L'une, de nets traits fréquentistes, l'autre épistémique. Pascal a remarquablement utilisé les principes de la dualité de la probabilité en situations diamétralement opposées. Le point commun fût l'incertitude, mais de caractéristiques différentes dans l'un et l'autre. Dans le problème de partage, l'incertitude est typiquement aléatoire, une incertitude ponctuelle irrésolue, un lancer, mais maîtrisée par une régularité sur le global, la série infinie. Dans le problème du pari, l'incertitude est d'origine épistémique, référée à nos connaissances.

La maîtrise sur le plan global de l'une permet d'établir des lois de probabilité, tandis que l'autre l'abordera par des arguments et des raisons devenant des principes acceptés par une communauté. Ces deux incertitudes nous conduisent non seulement à des notions de probabilités différentes, mais aussi à des procédés inférentiels particuliers, chacun d'eux conformant une école en Statistique.

Des éléments communs dans leurs manières de les évaluer

Ces deux notions de la probabilité arriveront par des voies distinctes à un point commun, l'ensemble fini de référence. La fréquentiste tire son "objectivité" des observations, mais cette démarche est très coûteuse et elle sera simplifiée par des moyens plus économiques que l'expérimentation. Lorsque les évidences permettent d'assumer l'équipossibilité des événements élémentaires, il suffira de compter les cas favorables parmi les cas possibles pour obtenir une estimation *"a priori"* des probabilités. Sous les conditions énoncées, cet ensemble permet d'anticiper la fréquence d'apparition d'un phénomène.

Les bayésiens arriveront par une autre voie à la même formule. L'ensemble de cas favorables sur cas possibles leur facilite à eux aussi la tâche d'évaluation d'une probabilité, sous le principe "naturel et logique" d'équipossibilité qui opère comme un motif à croire. De cette manière, Arnaud et Nicole évaluèrent dans la Logique de Port-Royal des probabilités sur la véracité ou fausseté d'un acte signé par des notaires.

Les deux probabilités trouvent donc un autre point en commun, non seulement nous avons un même terme probabilité et un même symbole, mais il y a aussi un même procédé : compter les cas favorables sur les cas possibles. Mais toujours pour évaluer des choses différentes, pour l'une on prévoit la fréquence d'apparition, pour l'autre on évalue combien on croît en la véracité d'une proposition.

Pour décrire cette situation nous avons introduit le terme *intrication*, en essayant de rendre compte de ce phénomène complexe de fusion sur le plan calculatoire de deux signifiés différents. Cette intrication entraînée par l'utilisation duale de la formule de Laplace en épreuves génériques contribue encore plus à la confusion entre les deux notions de la probabilité.

La Figure 8 nous montre les rôles de la probabilité fréquentiste et de l'ensemble de référence fini dans le procédé d'évaluation de la probabilité pour une épreuve générique. Cette figure pourrait bien décrire la représentation qu'un individu se ferait si sa perception des enjeux était sans ambiguïté : l'évaluation d'une épreuve générique où on fait recours à des ensembles de référence pour quantifier une certitude partielle.

Quelques procédés d'évaluation de probabilité bayésienne pour épreuves génériques

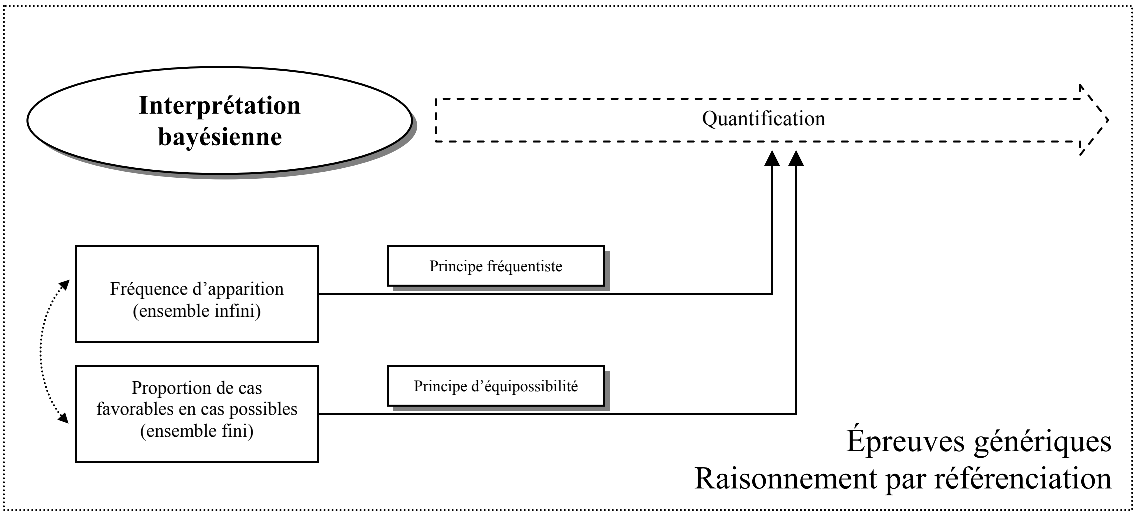


Figure 8

Mais en général, et particulièrement en classe de mathématiques, la pondération de ces éléments est toute autre. La relégation du signifié de la probabilité se voit compensée par une surestimation des aspects calculatoires amenant à une autre représentation de la situation où les interprétations se mêlent, se confondent ou même disparaissent. Le cas extrême, celui de la disparition, est bien représenté par ceux qui définissent la probabilité comme un quotient entre cas favorables et cas possibles. Les défenseurs de cette approche n'ont donné qu'une définition opératoire en rescindant les définitions sémantiques de la probabilité (Figure 9).

Définition de la probabilité opératoire pour épreuves génériques

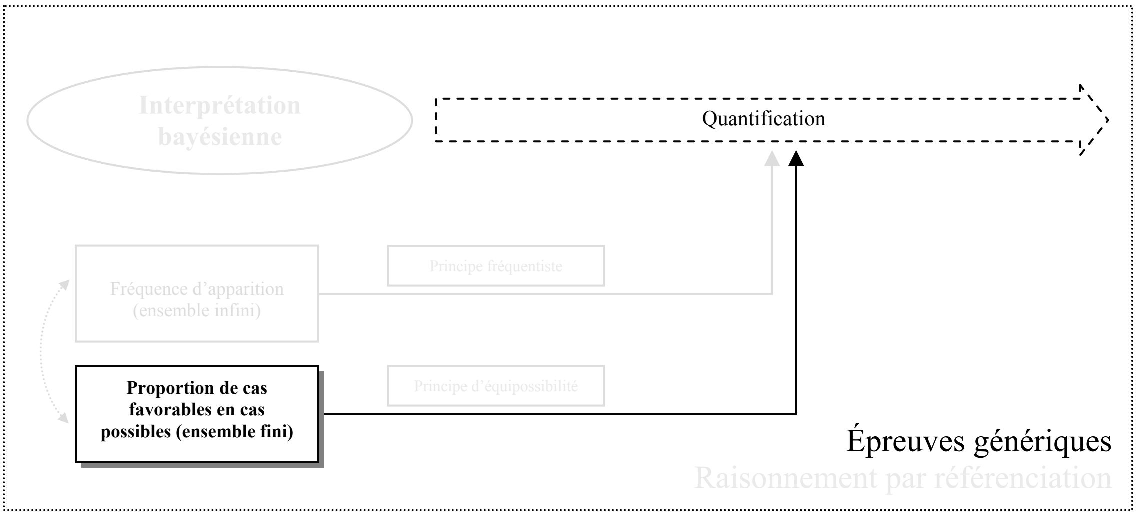


Figure 9

Une situation moins extrême est aussi usuelle où le rôle de la probabilité fréquentiste est retenu mais surestimé, en passant d'un motif à croire vers une position fictivement centrale dans le problème (Figure 10). Et de cette manière, en cachant l'une et en surlignant l'autre, on essaye de convaincre qu'il n'y a qu'une notion de probabilité, la fréquentiste.

Popper a indirectement mis en évidence qu'une épreuve générique est irrésoluble pour la probabilité fréquentiste. Il l'a fait par son intention de définir la probabilité comme une propension. Si probabiliser sur une épreuve générique avait été techniquement possible à partir de l'approche fréquentiste, ni Popper ni Peirce auraient eu la nécessité de concevoir une probabilité propensioniste.

Surestimation de la probabilité fréquentiste en épreuves génériques

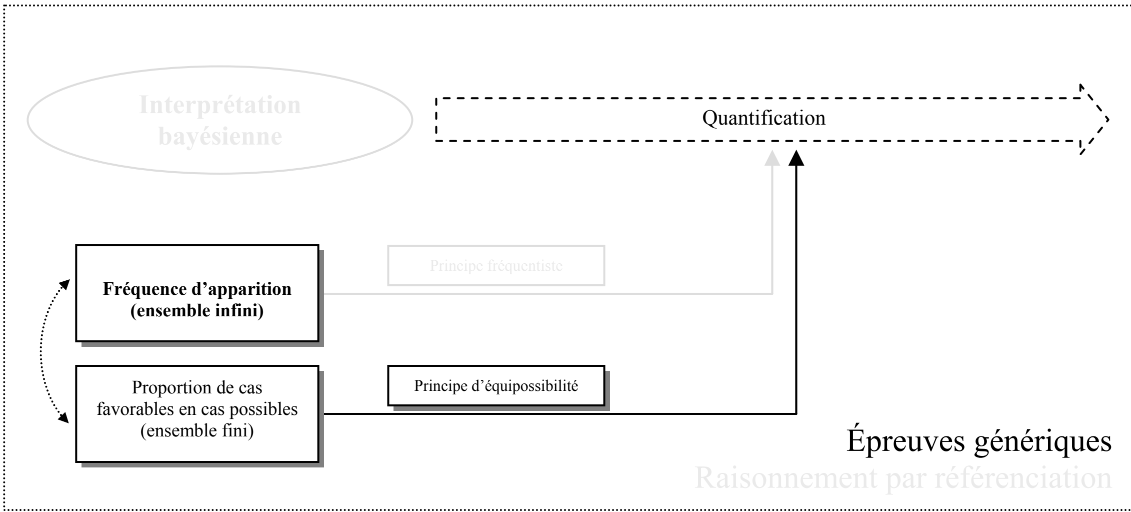


Figure 10

A l'utilisation duale de la formule de Laplace dans l'évaluation de la probabilité d'une épreuve générique, il faut donc joindre une autre source d'intrication, un autre ensemble de référence, cette fois-ci un ensemble infini, la fréquence d'apparition. En situations scolaires, les ensembles fini et infini se correspondent fréquemment l'un à l'autre, les extractions des urnes, la manipulation de cartes et les lancers de dés admettent l'un et immédiatement l'autre de ces ensembles. Ainsi l'évaluation de la probabilité d'une épreuve générique peut s'effectuer par ces deux ensembles usuellement interconnectés. Le cas le moins courant mais tout à fait valable, est celui d'une quantification de la probabilité d'une épreuve générique uniquement par la fréquence d'apparition du phénomène en question, le cas banal de probabiliser sur le prochain lancer de "ma" punaise en est un exemple. Il y a donc deux sources d'intrication (Figure 11), l'une est constituée par l'utilisation duale d'une même formule ; l'autre par la présence de la fréquence d'apparition dans l'évaluation d'une probabilité.

Intrication de procédés d'évaluation

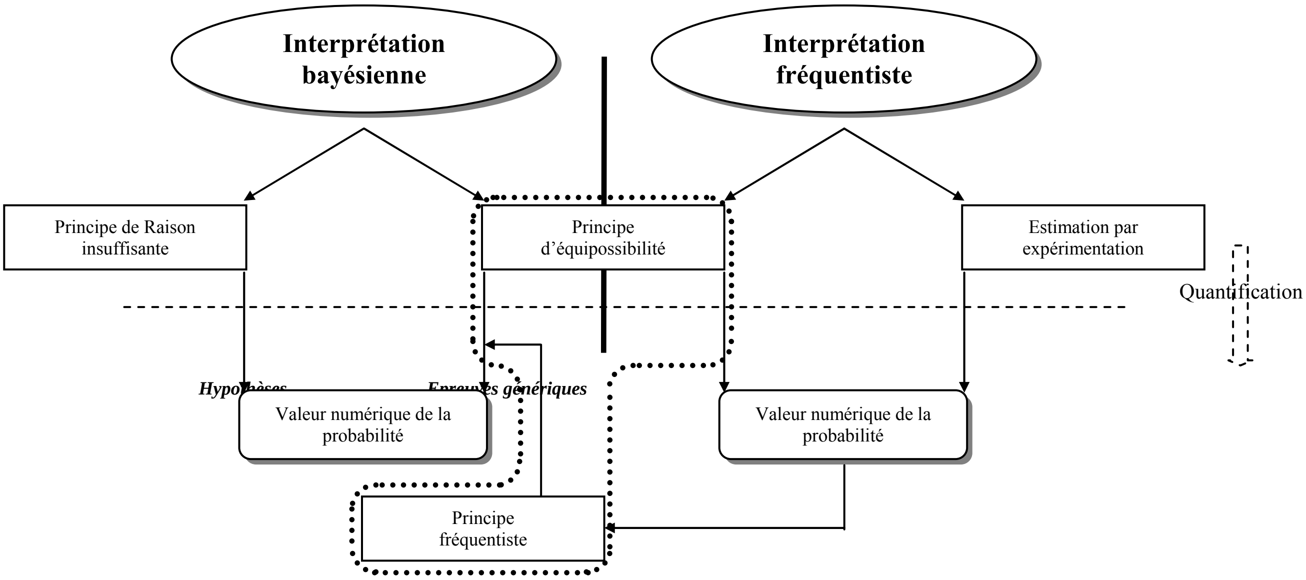


Figure 11

Les épreuves génériques sont donc des situations particulières mais très courantes où les deux probabilités participent, dans des rôles différents, mais intriquées de telle manière qu'il ne soit pas inhabituel de les confondre ou même de les fusionner en une seule probabilité, généralement la fréquentiste.

Un autre élément que nous souhaitons retenir de cette enquête concerne le caractère d'outil de la probabilité. Ce survol de l'histoire de la probabilité nous permet de décrire au moins très brièvement quelques conditions des contextes situationnels pour que la notion de probabilité devienne un concept nécessaire à la résolution du problème.

De son caractère d'outil

Ce qui nous intéresse ici est d'essayer de caractériser le genre de situations où on aurait besoin du concept *probabilité* et par conséquence le genre de situations qui demanderaient leur quantification, bien évidement dans ses deux acceptations.

Tout semble indiquer que l'évaluation ou quantification de la probabilité est une réponse à une nécessité de précision. En effet, la précision qu'apporte son évaluation numérique permet directement ou indirectement de trancher parmi un ensemble de possibilités. La manière directe impliquerait une simple comparaison de probabilités, l'indirecte, l'insertion de son évaluation dans des calculs plus complexes, tels que l'espérance mathématique.

D'un point de vue plus général, nous nous demandons quel genre de situations aurait besoin du concept *probabilité*. La probabilité émerge pour traiter l'incertain, il semblerait donc qu'un contexte indéterministe soit une condition pour justifier son recours. En d'autres termes, la probabilité serait un outil de prise de décision (au sens large) dans un contexte d'incertitude. Nous reviendrons sur ces conditions de fort intérêt didactique dans le *Chapitre III* consacré à nos expérimentations.

Des autres définitions de probabilité

Dans cette synthèse nous avons identifié plusieurs définitions ou interprétations de la probabilité. Les termes définition opératoire et définition sémantique nous ont permis d'en faire une première classification. La probabilité connue comme "classique" ou de "Laplace" définie comme un quotient de cas favorables sur cas possibles à été écartée par sa fonction exclusivement opératoire : nous avons traité ce sujet dans *Chapitre I. Quelle place pour la définition classique* et l'avons reconsidéré dans *Chapitre I. Des éléments communs dans leurs manières de les évaluer.*

D'autres notions ont été repérées, en particulier celles de probabilités logique et propensioniste. De sources bayésiennes, la première contribue à un mode de pierre angulaire de façon à ce que même si enrichie par l'approche d'une Logique inductive, elle ne devienne pas en soi une notion de probabilité qui s'impose en statistique. D'autre part, l'intention de Popper de substituer la notion fréquentiste par une de type propensioniste ne semble pas avoir prospéré; les traces de sa réussite ne sont pas évidentes à constater ; les approches inférentielles se concentrent toujours sur deux grandes écoles, la fréquentiste et la bayésienne.

D'une "dualité" vers une triade ?

A partir d'un point de vue statistique, la probabilité conçue juste comme un chiffre n'a pas d'intérêt. Cette seule composante de la probabilité ne permet pas de développer ni l'utilité de la probabilité pour une prise de décision dans ses formes même les plus élémentaires, ni la compréhension des procédés statistiques inférentielles non plus. Dans son rôle d'outil pour une prise de décision basée sur le critère de maximisation de l'espérance par exemple, l'habillage des calculs par leurs signifiés devient indispensable pour rendre compte des raisons de ce choix d'action.

Il y a donc deux axes, l'un concerne les interprétations, l'autre leurs évaluations numériques. En d'autres termes, un axe sémantique et l'autre opératoire. Le premier se réfère aux signifiés attribués aux termes, les deuxièmes aux formes de transformation sur le plan mathématique des probabilités. L'articulation entre ces axes s'effectue par un ensemble de principes acceptés au sein de chaque paradigme. Sans ces principes, les toutes premières évaluations de probabilités ne seraient qu'arbitraires et sans consensus. Ces principes conforment un cadre argumentatif pour le passage au nombre, rappelons par exemple la fonction du principe de raison insuffisante pour la probabilité bayésienne ou celle du principe d'équipossibilité pour la notion fréquentiste. Sans eux, il n'y aurait pas de raisons pour leurs évaluations numériques.

La Figure 12 résume l'architecture de la Figure 11. Elle représente l'enchaînement de trois aspects de la probabilité : dans un extrême les interprétations, dans l'autre les calculs, entre eux, un ensemble de principes permettant le passage au nombre. Ces principes ont attiré l'attention des savants des deux approches de la probabilité. Cette triade sera d'intérêt didactique, nous reviendrons plus tard (*Chapitre III*) sur ce sujet lors de nos expérimentations.

Triade Interprétations –principes – évaluation numérique de la probabilité

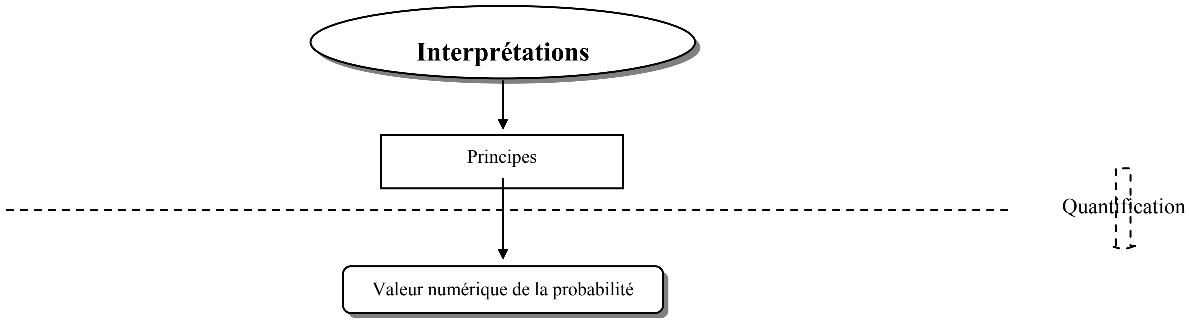


Figure 12

L'autre objectif principal pour ce premier chapitre concernait le repérage d'éléments caractéristiques à chacune des interprétations de la probabilité. Nous en avons identifié quatre, nous résumerons les idées principales de ces éléments caractéristiques.

Des éléments caractéristiques

Les objets sur lesquels chaque approche peut probabiliser

Le terme "objet" a été introduit pour représenter l'argument A dans l'expression P(A). Notre intention est d'éviter l'usage généralisé du terme "événement", plutôt associé à un contexte fréquentiste. Ceci étant, chacune de ces approches admet des objets de nature différente dans leurs respectives interprétations de la probabilité.

Nous avons identifié trois catégories relativement disjointes. Deux de ces catégories pour la probabilité bayésienne, une troisième pour la fréquentiste. Un groupe d'objets pour la probabilité bayésienne est celui des épreuves génériques, l'autre celui des hypothèses. Pour la probabilité fréquentiste, il y a une catégorie unique regroupant les séries infinies.

La notion de série infinie est équivalente à celle de *collective* de Richard von Mises. Une série infinie serait donc caractérisée par deux clauses : la première traite sur la convergence d'apparition d'un phénomène lors de la répétition jusqu'à l'infini de l'épreuve sous les mêmes conditions, la deuxième sur la randomisation de la série. L'interprétation fréquentiste probabilise donc sur la série infinie d'occurrences d'un phénomène ou événement donné. Ceci constitue donc le seul objet sur lequel il est possible de probabiliser pour cette approche.

Pour la probabilité bayésienne nous avons différencié deux types d'objets. L'un concerne les épreuves génériques, l'autre les hypothèses. Les épreuves génériques sont caractérisées par sa plausibilité de reproduction sous des conditions identiques. La principale différence avec la série infinie (ou *collective* de von Mises) réside en ce que l'épreuve générique est un élément de la série infinie. Par ailleurs, son caractère générique est donné par le manque d'éléments identificateurs permettant d'individualiser l'épreuve, elle est donc perçue comme une parmi d'autres, sans aucune distinction des épreuves plausibles de se réaliser.

L'autre ensemble d'objets pour la probabilité bayésienne est celui des hypothèses. Ce sont des propositions de la même manière que les épreuves génériques mais les conditions du problème empêchent de se représenter la reproduction de l'épreuve, la perception de son caractère épistémique est plus évidente pour ces objets par la non interférence de la composante "aléatoire" présente dans les épreuves génériques.

Une deuxième différence est celle des types de raisonnements mobilisés par chacune des interprétations, cette deuxième différence est en étroit rapport avec celle des valeurs logiques.

Les types de raisonnements

La probabilité fréquentiste est une caractéristique d'une série infinie, pour autant, elle est théorique. Cette approche a besoin de la construction d'un modèle pour estimer la valeur de la probabilité. Par leurs caractéristiques théoriques, les probabilités sont donc estimées à partir d'informations ou d'expériences constatées. Une fois les paramètres du modèle acceptés, la détermination de la probabilité se réalise par déduction. En admettant les hypothèses, il est possible de conclure sur la valeur de la fréquence d'apparition à long terme d'un événement donné.

Pour le paradigme bayésien, deux types de raisonnements sont mobilisés. Pour les épreuves génériques, nous avons introduit le terme de raisonnement par référenciation. Ce procédé argumentatif consiste à considérer des ensembles de référence en mode de motifs à croire. Nous avons repérés deux types d'ensembles, l'ensemble fini de référence qui conduit à l'utilisation de la formule de Laplace et l'ensemble infini de référence constitué par la fréquence apparition du phénomène en question. Un deuxième type de raisonnement est possible pour la probabilité bayésienne. Il s'agit de l'abduction ou recherche de la meilleure explication possible. Cette démarche est mise en place lorsque l'on probabilise sur un ensemble d'hypothèses en considérant l'information disponible. La représentation symbolique de cette expression est la probabilité conditionnelle. A ces types de raisonnements correspondent des valeurs logiques et des principes d'évaluations propres. Nous commencerons par les valeurs logiques associées aux respectives interprétations de la probabilité.

Les valeurs logiques

Les hypothèses pour le paradigme fréquentiste admises, les valeurs logiques associées à chacune des épreuves de la série sont ceux de la dichotomie vrai/faux, mais ces valeurs ne sont qu'attribuables après coup, après avoir renversé l'incertitude propre à l'utilisation de la probabilité. La valeur logique portée sur la série est par définition aussi vraie (ou faux).

Pour la notion bayésienne, l'espace de valeurs possibles devient continu, allant de zéro à un, les deux extrêmes représentant deux certitudes. Par ailleurs, si pour la probabilité fréquentiste la valeur de la probabilité est fixe et propre à une série (ou *collective* d'après von Mises), l'approche bayésienne elle, peut évoluer en fonction de l'arrivée des informations du système. Ceci a été le cas pour l'exemple de la bouteille où chaque retournement donnait des informations permettant de reconsidérer les valeurs des certitudes en fonctions des couleurs de billes apparues à travers le bouchon transparent.

Les principes mobilisés

Cette catégorie traite des principes sur lesquels se fondent les évaluations des probabilités. Ces principes argumentent le passage au nombre des probabilités (Figure 7 et 8). Sans eux, il n'y aurait pas de raisons pour l'utilisation, entre autres, de la formule de Laplace. Sans eux, les évaluations des probabilités bayésiennes seraient non seulement subjectives mais livrées aux choix personnels. Bien qu'un tel choix soit possible pour un subjectiviste comme Bruno de Finetti, il nous semble nécessaire de considérer ces principes à partir d'une intention didactique comme la nôtre. En effet, proposer l'approche "extrême" de De Finetti à un professeur où chaque élève évaluerait à son choix les probabilités pourrait être trop risquée par ses difficultés en l'institutionnalisation, en l'évaluation et même par les attentes liées au contrat didactique en vigueur. Nous reviendrons sur ce sujet dans le *Chapitre III* consacré aux expérimentations.

Pour ce qui concerne l'approche fréquentiste et en amont d'une estimation par des épreuves répétées, le moyen pour évaluer la probabilité est celui du principe d'équipossibilité qui permet lui aussi une estimation mais par l'inspection de la configuration du système qui produit les épreuves. Ainsi, si une hypothèse d'équipossibilité d'événements élémentaires est admissible, on peut déduire une probabilité par combinaison des cas élémentaires.

L'approche bayésienne se servira aussi de ce principe pour évaluer la probabilité d'une épreuve générique où la référence est un ensemble fini. Si par contre on ne dispose que de la fréquence d'apparition du phénomène étudié, l'évaluation de la probabilité se ferait par le principe fréquentiste. Finalement, pour les hypothèses, le principe retenu est celui de raison insuffisante applicable lors de situations où l'on ne dispose d'aucune information autre que l'ensemble d'hypothèses possibles. Le Tableau 5 est une reproduction du Tableau 4 qui synthétise les quatre éléments caractéristiques des deux interprétations de la probabilité.

Quatre différences entre les interprétations de la probabilité

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="5">Interprétations</td></tr>
<tr><td colspan="2">Fréquentiste</td><td colspan="3">Bayésienne</td></tr>
<tr><td rowspan="4">Différences</td><td>Objet sur le quel probabiliser</td><td colspan="2">Série infinie</td><td colspan="2">Epreuve générique</td><td>Hypothèse</td></tr>
<tr><td>Valeurs logiques portés sur ces objets</td><td colspan="2">Vrai- faux</td><td colspan="3">Intervalle continue [0,1]</td></tr>
<tr><td>Type de raisonnement</td><td colspan="2">Déduction</td><td colspan="2">Référenciation</td><td>Abduction</td></tr>
<tr><td>Quelques principes de quantification</td><td>Hypothèses du modèle</td><td colspan="2">Equipossibilité (Ensemble fini)</td><td>Fréquentiste (ensemble infini)</td><td>Raison insuffisante</td></tr>
</table>

Tableau 5

Ces éléments caractéristiques nous permettrons d'interroger le profil des exercices du chapitre probabilité des manuels de lycée français, c'est le sujet du *Chapitre II* que nous développerons ci-dessous.

Chapitre II : Analyse de manuels

2.1 Introduction

Dans le premier chapitre nous avons repéré des traces de la dualité d'interprétation de la probabilité. Notre liste d'auteurs consultés a commencé par Blaise Pascal qui dans ses lettres adressées à Fermat a proposé une solution à un problème de partage où la pierre angulaire fut la notion fréquentiste. Le même auteur dans *Pensées* (Pascal, 1670) s'est occupé d'un problème radicalement différent connu comme "le grand pari" : la solution de Pascal s'est appuyée sur les éléments de la théorie de la décision, l'argument ayant été la maximisation de l'espérance par la dominance en faveur de la croyance en Dieu contre une position incrédule de son existence.

Ainsi Pascal fut un des premiers à transférer un raisonnement sur le hasard à une inférence ne reposant pas sur un processus aléatoire. Leibniz a repris cette problématique en insistant sur la nécessité de construire une nouvelle sorte de logique pour traiter l'incertain. Keynes et Carnap ont approfondis cette approche en proposant une notion de probabilité dans la perspective d'une *Logique inductive*.

La notion fréquentiste pour sa part s'est aussi consolidée, conservant toujours son lien avec la notion bayésienne. Bref, une dualité d'interprétation qui jusqu'à aujourd'hui se manifeste par deux approches inférentielles largement confirmées dans les sociétés savantes.

Notre intention pour le premier chapitre fut de montrer l'existence d'une probabilité à deux visages. Ce deuxième chapitre sera consacré à rechercher cette dualité dans les manuels de Première année de lycée en France. En d'autres termes, nous nous intéresserons au processus de transposition didactique de la dualité de la probabilité en observant son éventuelle présence dans les manuels français.

Entre la dualité de la probabilité en tant que référence épistémologique et l'objet transposé dans les manuels se trouvent les documents officiels. Nous allons nous intéresser aussi, très brièvement, à l'approche choisie par les rédacteurs des programmes par rapport à la dualité de la probabilité.

Nous verrons dans programmes de l'année 2001 que le choix de ses rédacteurs s'incline vers une seule de ces notions, la fréquentiste, et qu'en consonance avec ces documents normatifs, les manuels de lycée présentent la probabilité à partir de cette approche uniquement. Nous essayerons de voir si dans les faits, un découpage de cette notion est possible ou si au contraire, la dualité d'interprétation se manifeste dans les exercices des manuels, et cela en amont des directives officielles.

Nous avançons une des principales hypothèses de ce chapitre : nous postulons que la dualité est une caractéristique intrinsèque à la probabilité, donc incontournable ; et qu'une éventuelle tentative de découpage n'est possible qu'en glissant (Carranza, 2005) l'intérêt vers un travail mathématique au détriment du signifié statistique de la notion de probabilité.

Nous pouvons affirmer qu'un de nos principaux objectifs pour ce chapitre consiste à confirmer la condition d'inévitabilité de la dualité. Nous essayerons de confirmer cette hypothèse par la présence de situations fréquentistes et bayésiennes dans les exercices des manuels de lycée français. Pour cela nous analyserons les exercices du chapitre Probabilité de manuels de deux filières de lycée, Première ES et Première S.

Nous avons choisi la classe de Première parce c'est cette année que la probabilité acquière le statut d'objet à enseigner ; et les filières ES et S pour l'espace relativement important consacré à cette notion dans ces sections, au moins par rapport à d'autres filières.

Cette étude sur les manuels pourrait se formaliser par la proposition d'hypothèses à tester sur l'ensemble des exercices. Des hypothèses qui s'exprimeraient sous la forme d'implications telles que si *a alors b.* Néanmoins, la variabilité propre à l'ensemble des exercices fait qu'une possible validation ne soit que partielle (si *a alors presque b).* Par exemple, une des hypothèses pourrait s'exprimer sous la forme *si le contexte d'un exercice est bayésien alors son interprétation ne sera pas traitée »* ou *« si un exercice est une épreuve générique alors le rôle de l'interprétation fréquentiste sera surévalué »,* etc. etc. Nous analyserons les exercices sous une perspective statistique, en prenant un ensemble de variables. Nous étudierons les tendances sous la forme de quasi implications des exercices de quelques manuels français.

Nous nous servirons de la théorie de l'Analyse statistique implicative (Gras et al., 1996) et du logiciel CHIC pour traiter statistiquement notre hypothèse d'incontournable dualité de la probabilité. Bien évidement, cette étude servira aussi pour explorer d'autres tendances et relations présentes dans les exercices.

Premièrement nous présenterons donc une synthèse de l'approche des rédacteurs des programmes, puis nous aborderons en détail notre étude sur les manuels de la classe de Première ES et S.

2.2 L'approche des rédacteurs des programmes

Les manuels analysés correspondent aux programmes publiés le 31 août de l'année 2000 dans le Bulletin officiel (B.O.) du Ministère de l'Education nationale et de la Recherche, ils sont téléchargeables sur le site web du Ministère (www.education.gouv.fr, voir *Annexe*).

Pour la classe de Première série Scientifique enseignement obligatoire, le nouveau programme applicable à compter de l'année scolaire 2001-2002 se trouve dans H.S. N° 7. Pour la classe de Première série Economique et Sociale dans le H.S. N° 8. Nous transcrivons pour chaque filière les idées principales concernant l'approche de l'interprétation de la probabilité dans Programmes et les Documents d'accompagnement.

Filière Economique et Sociale

Le programme se trouve dans l'Annexe 2 du B.O.. Il comporte 4 sections : *Objectifs généraux pour la série ES*, *Mathématiques et informatique en Première ES*, *Organisation de l'enseignement et du travail des élèves* et *Les contenus du programme de la classe de Première ES*. Suite à cette présentation, les contenus sont affichés sous la forme d'un tableau à la page 7 du H.S. N° 8, ci-contre nous reproduisons la partie consacrée au chapitre Probabilités de ce tableau et une partie du document d'accompagnement :

Fragment du Programme de mathématiques. Première ES. Probabilités

Probabilités Définition d'une loi de probabilité sur un ensemble fini. Probabilité d'un événement, de la réunion et de l'intersection d'événements. Modélisation d'expériences de référence menant à l'équiprobabilité; utilisation de modèles définis à partir de fréquences observées.	Le lien entre loi de probabilité et distribution de fréquences sera éclairé par un énoncé vulgarisé de la loi des grands nombres. On mènera de pair simulation et étude théorique de la somme de deux dés (en liaison avec le paragraphe précédent).	Un énoncé vulgarisé de la loi des grands nombres peut être par exemple: *Pour une expérience donnée, dans le modèle défini par une loi de probabilité P, les distributions des fréquences obtenues sur des séries de taille n se rapprochent de P quand n devient grand.* On indiquera que simuler une expérience consiste à simuler un modèle de cette expérience. On pourra ne pas se limiter à l'étude d'une seule situation et envisager d'autres expériences (produit de deux dés, somme de trois dés…). On pourra repérer les difficultés soulevées par le choix d'un modèle mais sans s'y attarder: on utilisera directement des modèles que la statistique a permis de choisir.

Fragment du Document d'accompagnement. Première ES page 15

Modélisation à partir de fréquences

On se contentera, pour certains exercices, de fournir un modèle en indiquant dans un premier temps que des techniques statistiques ont permis de le déterminer et de le valider à partir de nombreuses données expérimentales. Pour déterminer et/ou valider un modèle probabiliste, le premier outil dont on dispose est un théorème de mathématiques appelé « loi des grands nombres », dont un énoncé intuitif est :

Dans le monde théorique défini par une loi de probabilité P sur un ensemble E, les fréquences des éléments de E dans une suite de n expériences identiques et indépendantes tendent vers leur probabilité quand n augmente indéfiniment.

Ou encore :

Si on choisit n éléments selon une loi de probabilité P, indépendamment les uns des autres, alors la distribution des fréquences est voisine de P lorsque n est grand.

Filière Scientifique

Le programme se trouve dans l'Annexe 2 du B.O.. Il comporte 5 sections : *Généralités à propos d'une formation scientifique en première et en terminale S*, *Mathématiques et informatique en terminale S*, *Epistémologie et histoire des mathématiques, Organisation de l'enseignement et du travail des élèves* et *Les contenus du programme de la classe de Première ES*. Suite à cette présentation, les contenus sont affichés sous la forme d'un tableau à partir la page 167 du H.S. N° 7, ci-contre nous reproduisons la partie consacrée au chapitre Probabilités de ce tableau et une partie du document d'accompagnement :

Fragment du Programme de mathématiques. Première S. Probabilités

Probabilités Définition d'une loi de probabilité sur un ensemble fini. Espérance, variance, écart-type d'une loi de probabilité. Probabilité d'un événement, de la réunion et de l'intersection d'événements. Cas de l'équiprobabilité. Variable aléatoire, loi d'une variable aléatoire, espérance, variance, écart-type. Modélisation d'expériences aléatoires de référence (lancers d'un ou plusieurs dés ou pièces discernables ou non, tirage au hasard dans une urne, choix de chiffres au hasard, etc.).	Le lien entre loi de probabilité et distributions de fréquences sera éclairé par un énoncé vulgarisé de la loi des grands nombres. On expliquera ainsi la convergence des moyennes vers l'espérance et des variances empiriques vers les variances théoriques; on illustrera ceci par des simulations dans des cas simples. On pourra aussi illustrer cette loi avec les diagrammes en boîtes obtenus en simulant par exemple 100 sondages de taille *n*, pour *n* = 10; 100; 1000. On simulera des lois de probabilités simples obtenues comme images d'une loi équirépartie par une variable aléatoire (sondage, somme des faces de deux dés, etc.).	On pourra par exemple choisir comme énoncé vulgarisé de la loi des grands nombres la proposition suivante: *Pour une expérience donnée, dans le modèle défini par une loi de probabilité P, les distributions des fréquences calculées sur des séries de taille n se rapprochent de P quand n devient grand.* On indiquera que simuler une expérience consiste à simuler un modèle de cette expérience. La modélisation avec des lois ne découlant pas d'une loi équirépartie est hors programme. On évitera le calcul systématique et sans but précis de l'espérance et de la variance de lois de probabilité.

Fragment du Document d'accompagnement. Première S page 71

Pour valider un modèle probabiliste, le premier outil dont on dispose est un **théorème de mathématiques** appelé « loi (forte) des grands nombres », dont un énoncé intuitif est:
Dans le monde théorique défini par une loi de probabilité P sur un ensemble fini, les fréquences des éléments de cet ensemble dans une suite de n expériences identiques et indépendantes « tendent » vers leur probabilité quand n augmente indéfiniment.
Ou encore :
Si on choisit n éléments selon une loi de probabilité P, indépendamment les uns des autres, alors la distribution des fréquences est voisine de P lorsque n est grand.

Bref, les deux programmes et leurs respectifs documents d'accompagnement définissent explicitement la probabilité comme la stabilisation de fréquences, aucune mention n'est faite à la notion bayésienne de la probabilité. D'ailleurs, les simulations occupent une place centrale pour les rédacteurs de ces programmes.

Enfin, il y a donc une transformation significative de la probabilité dans cette étape de la transposition didactique, les programmes ne retenant qu'un de ses visages de la dualité, l'approche fréquentiste. Nous analyserons si cette intention fait écho dans les exercices proposés par les manuels de ces deux filières.

2.3 Problématique

Les programmes officiels en tant que documents normatifs établissent que la notion de probabilité s'associe (ou doit s'associer) à la stabilisation de fréquences. Nous nous intéresserons aux manuels en tant qu'instances suivant les programmes officiels dans le processus de transposition didactique de la notion de probabilité.

Dans notre étude de manuels, nous n'analyserons pas les possibles marges de manœuvre que les rédacteurs pourraient avoir face aux directives des programmes, ni non plus des possibles interprétations personnelles plus ou moins éloignées de celles des rédacteurs des documents officiels. Nous admettrons par la suite qu'il y a une relative assimilation et appropriation de la directive ministérielle de considérer la probabilité comme étant exclusivement fréquentiste et que cette consigne ministérielle devienne, d'une certaine manière, la leur au moment de la rédaction de manuels.

Nous nous intéresserons à une analyse des exercices en termes d'espace potentiel de travail (Kuzniak, 2004). Les exercices seront considérés par leurs caractéristiques contextuelles, ces dernières nous permettront de reconnaître la présence potentielle d'une ou l'autre des interprétations. C'est dans ce sens que nous parlons d'un exercice comme un espace potentiel de travail, un lieu qui par ses caractéristiques favoriserait l'évocation de l'interprétation fréquentiste ou la bayésienne. Ces espaces n'étant que potentiels, nous ignorons de quelle manière professeurs et élèves aboutissent dans les faits à traiter ces exercices.

Les caractéristiques que nous observerons dans les exercices sont liées aux éléments caractéristiques que nous avons présentés dans le premier chapitre. Ces éléments contribueront à déterminer si l'espace potentiel d'un exercice est fréquentiste, bayésien ou ambigu. Nous envisageons donc de confirmer notre hypothèse de l'incontournable dualité d'interprétation de la probabilité par la présence de ces différents contextes dans les exercices des manuels. En effet, c'est en verifiant l'existence de problèmes répondant au paradigme bayésien et au fréquentiste que nous confirmerons notre principale hypothèse pour ce chapitre, l'incontournable dualité de la probabilité.

Une autre question sera traitée dans cette analyse, si notre hypothèse d'incontournable dualité est vérifiée : nous tenterons de déterminer les moyens mis en place pour résoudre l'apparente contradiction provoquée par l'intention d'institutionnaliser une seule notion lorsque dans les faits les deux sont présentes dans les exercices.

Finalement, cette étude nous permettra aussi d'estimer le genre de situations auxquelles auraient été confrontés les élèves de BTS dans leur passage pour le lycée. En principe les manuels de l'année 2001 étaient ceux utilisés par les élèves de BTS de l'année 2007 lors de leur passage par la classe de Première au lycée. Nous commencerons par une brève description des manuels concernés.

Les manuels observés

Pour ce deuxième chapitre nous ne prétendons pas faire une étude exhaustive des manuels disponibles en France métropolitaine, ni non plus valider la représentativité de nos inférences par des méthodes statistiques. Dans notre étude, nous nous contenterons d'analyser exhaustivement les exercices du chapitre Probabilité de quatre manuels de la classe de Première de lycée français, deux correspondants à la filière Economie et sciences sociales, les deux autres à la filière Scientifique.

Bien qu'à partir de cette analyse aucune inférence ne peut être méthodologiquement validée, nous admettrons une relative homogénéité entre manuels de la même filière due principalement à la contrainte qui pèse sur eux de répondre aux directives émanant du Ministère de l'Education Nationale.

Nous baserons les quelques inférences tirées des analyses des quatre manuels sur cette supposée relative homogénéité entre manuels de la même filière. Ces manuels correspondent à quatre maisons d'édition françaises, pour la filière Economie et sciences sociales : Bréal et Didier ; pour la filière Scientifique, Nathan et Belin.

Notre étude ne portera que sur les exercices du chapitre Probabilité, principalement parce que notre hypothèse établit que la dualité devient incontournable dans les faits, au moment de la résolution de situations-problèmes. Et cela en amont de toute tentative déclarative (y compris éventuellement dans la partie Cours des manuels) de la présenter comme étant uniquement fréquentiste. Nous préciserons quelques détails de ces manuels correspondants au programme de l'année 2001.

Filière Economie et Sciences sociales

Bréal. Mathématiques 1ère ES enseignement obligatoire et option (Bréal ES)

Ce manuel de 287 pages fut édité en mai 2001. Le chapitre Probabilité est placé dans le deuxième quart du manuel, plus précisément entre les pages 84 et 104, les exercices commençant à la page 98. Cette partie du chapitre compte donc avec sept pages avec un total de trente et un exercices repartis en deux sections, la première appelée *Maîtriser le cours* (Exercices 1 au 12), la deuxième *Savoir-faire fondamentaux* (Exercices 13 au 31).

Didier. Mathématiques Obligatoire Première ES. Collection Dimathème (Didier ES)

La maison Didier publie ce manuel de 222 pages en juin 2001. Le chapitre Probabilité se trouve dans le dernier quart, dès la page 187 jusqu'à la 213. La partie consacrée aux exercices comporte trois sections, *Travail autonome* (Exercices 1 au 5), *Pour s'entraîner* (Exercices 6 au 50) et *Pour chercher* (Exercices 51 au 65).

Filière Scientifique

Belin. Mathématiques Première S (Belin S)

De la maison Belin, ce manuel paru en janvier 2001 comporte 290 pages. Le chapitre Probabilité est placé dans le dernier quart, dès la page 194 à la 223. Les exercices commencent à partir de la page 206. Deux grandes sections sont identifiées, la première des *Exercices* (Exercices 1 au 92) et *Problèmes* (Exercices 93 au 104).

Nathan. Mathématiques Première S. Collection Transmath (Nathan S)

Ce dernier manuel, édité par la maison Nathan, de 432 pages a été publié en mai 2001. Le chapitre Probabilité commence à la page 205 et termine à la 234. Les exercices commencent à partir de la page 224 et sont répartis en quatre sections : *Maîtriser le cours* (Exercices 1 au 20) Pour apprendre à chercher (Exercices 21 au 25), *Pour progresser* (Exercices 26 au 76) et *Problèmes de synthèse* (Exercices 68 au 73).

2.4 Méthodologie

Pour tester nos hypothèses nous avons considéré un ensemble de variables binaires. Ces variables déclinées en modalités assument la valeur 1 lorsqu'un exercice répond à la modalité en question, et 0 lors du cas contraire. Pour toutes nos variables, leurs modalités constituent une partition.

Le Tableau 6 résume le nombre d'exercices par manuel. Chacun de ses manuels sera statistiquement traité de manière individuelle. Le nombre relativement faible d'exercices chez certains manuels nous a mené à faire des regroupements de modalités, en particulier pour ceux de la filière Economie et sciences sociales. Les regroupements ont visé un nombre minimum de cinq effectifs par modalité. Néanmoins, quelques-unes n'ont pas été regroupées, précisément pour l'intérêt qui soulève son faible nombre d'effectifs. Nous décrirons les variables retenues et leurs respectives modalités.

Nombre d'exercices par manuel

Manuel	Nombre d'exercices
BELIN 1S	*104*
NATHAN 1S	*73*
BREAL ES	*31*
DIDIER ES	*65*

Tableau 6

Les variables et première lecture

Un premier groupe appelé *variables générales* cherche entre autres, à décrire la longueur des exercices. Dans le premier chapitre nous avons commenté que la dualité de la probabilité ne se manifeste ni dans les registres symbolique, ni numérique et qu'en plus un même terme (probabilité) est employé comme signifiant dans le registre langagier pour les deux interprétations. Différencier les interprétations requerrait en principe un travail important dans le registre de la langue naturelle (écrite dans le cas des exercices des manuels). Ce registre serait le seul qui par sa richesse permettrait de développer les différences entre ces deux interprétations.

D'ailleurs, toute tâche concernant les interprétations demande un travail additionnel à celui des calculs, un travail qui en principe se verrait manifesté par des exercices relativement longs, avec un retour à la fin sur le registre langagier pour interroger sur l'interprétation des calculs réalisés. Ce premier groupe de variables tend à repérer cette caractéristique en principe nécessaire au développement d'une interprétation de la probabilité, des exercices longs.

Dans ce groupe de variables, nous y en incluons d'autres telles que l'utilisation de registres autres que celui de la langue, les principales tâches mathématiques involucrées dans leur résolution et le genre de contexte dans lequel se déroule la situation-problème.

Un deuxième groupe de variables tend à préciser l'interprétation sous-jacente dans chaque exercice et à identifier dans le cadre d'une dialectique outil-objet (Douady, 1986), le rôle de la probabilité en tant qu'outil, cette fois-ci, outil d'aide à la prise de décision dans un contexte d'incertitude.

Nous décrirons les variables retenues et leurs respectives modalités. Nous présenterons une synthèse des distributions des exercices des quatre manuels pour chacune de ces variables. Cette étude descriptive sera complétée par une autre de type inférentiel tendant à valider nos hypothèses de l'incontournable dualité d'interprétation de la probabilité.

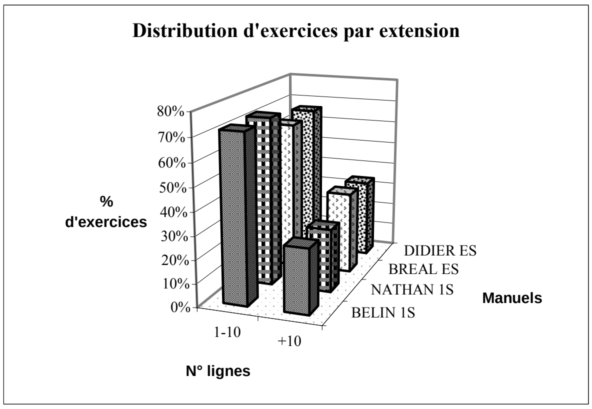
Variables Générales

Nombre de Lignes

Pour les manuels concernés, les exercices sont présentés en deux colonnes par page. Nous avons décidé de retenir le nombre de lignes des textes plus ou moins complets. Quatre modalités avaient été considérées en principe ([1 à 10] ; [11 à 20] ; … ; [31 à 40]) mais le faible nombre d'effectifs dans les dernières nous a obligé à reconsidérer ce découpage ; nous en avons finalement retenu deux modalités ([1 à 10] et [+ de 10]).

Le Graphique 5 représente le pourcentage d'exercices par nombre de lignes de texte et par manuel. Il y a en général, une tendance à proposer des exercices courts, la proportion de ces exercices varie entre 64,5 % (Bréal ES, 20/31 exercices) et 72,1% (Belin S, 75/104 exercices) (Tableau 7).

Nous constatons une proximité entre manuels de la même filière que nous ne validerons pas statistiquement. D'ailleurs, le nombre d'exercices proposés par les manuels ES a tendance à être inférieur à ceux de la filière S (104 et 73 pour S ; 65 et 31 pour ES). Ces derniers laissent plus de place en termes de pourcentage aux exercices longs (33,8% et 35,5% pour ES ; 27,4% et 27,9% pour S).



Graphique 5

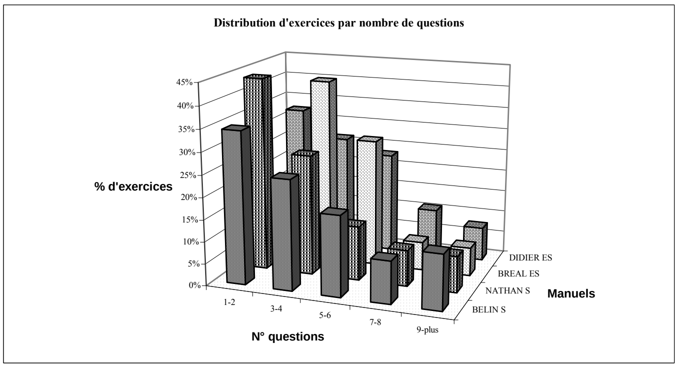
Distribution d’exercices par nombre de lignes de texte (effectifs et pourcentage)

Manuel	Nombre de Lignes 1-10		+10		Total
BELIN S	*75*	*72,1%*	*29*	*27,9%*	104
NATHAN S	*53*	*72,6%*	*20*	*27,4%*	73
BREAL ES	*20*	*64,5%*	*11*	*35,5%*	31
DIDIER ES	*43*	*66,2%*	*22*	*33,8%*	65

Tableau 7

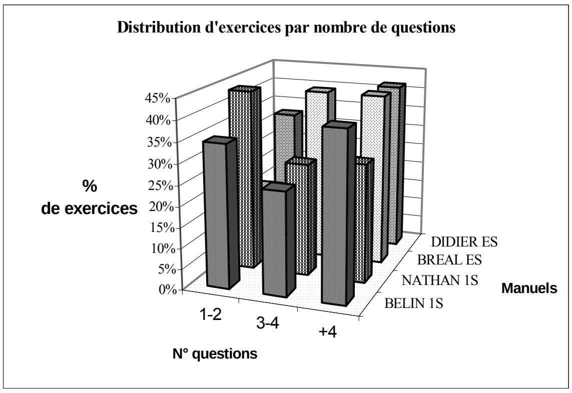
Nombre de Questions

Cette autre variable vise repérer le nombre de questions ou de tâches proposées dans un exercice. Originellement, les modalités était cinq ([1 à 2] ; [3 à 4] ; [5 à 6] ; [7 à 8] ; [+ 8]) (Graphique 6). La distribution des exercices nous montre une tendance à poser un nombre réduit de questions, en particulier le manuel de la maison Nathan (filière Scientifique).



Graphique 6

Pour des raisons toujours d’effectifs nous avons procédé à un regroupement des trois dernières modalités (Graphique 7). Dans cette nouvelle distribution, la dernière (plus de quatre questions) est très large et concentre plus de 40% des exercices pour trois manuels (Belin S, Breal ES et Didier ES). Bien que plus informative configuration du Graphique 6, nous retiendrons celle du Graphique 7.



Graphique 7

Distribution d'exercices par nombre de questions (effectifs et pourcentage)

Manuel	Nombre de questions 1-2		3-4		+4		Total
BELIN S	*36*	*34,6%*	*26*	*25,0%*	*42*	*40,4%*	104
NATHAN S	*32*	*43,8%*	*20*	*27,4%*	*21*	*28,8%*	73
BREAL ES	*5*	*16,1%*	*13*	*41,9%*	*13*	*41,9%*	31
DIDIER ES	*21*	*32,3%*	*17*	*26,2%*	*27*	*41,5%*	65

Tableau 8

Quelques types de représentations proposées

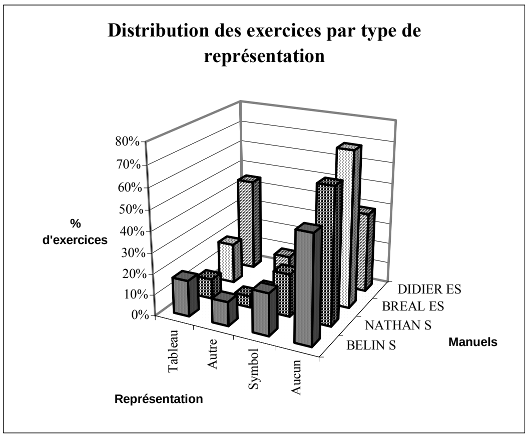
Avec cette variable, nous nous intéressons à repérer l'utilisation d'autres modes de représentation que les registres symbolique ou numérique. Les représentations graphiques favoriseraient la mise en relation et l'intégration de l'information ainsi que la lecture interprétative et l'exploration de relations cachées entre les données (tel que le diagramme de Pareto par exemple), en opposition aux registres symboliques et numériques qui faciliteraient en principe les transformations numériques.

Le Graphique 8 représente la distribution en termes de pourcentage de la présence de différents types de représentation autre que le langage. Ce graphique montre une centration sur deux types de représentations (*Tableau* et *Symbolique*), il met aussi en évidence le fort taux d'exercices sans référence explicite à d'autres types de représentations que le langagier.

Signalons qu'aucun histogramme n'a été proposé dans ces manuels et que le pourcentage relativement faible de représentation symbolique est dû à notre choix de retenir l'autre type de représentation si le symbolique n'était pas le seul à se manifester. Parmi les combinaisons possibles de registres, le couple *Symbolique/Tableau* est le plus fréquent, dans ces cas, nous n'avons retenu que la représentation *Tableau*.

Dans les manuels observés, le tableau semble avoir la fonction d'organiser la présentation des probabilités dans le registre numérique. La représentation symbolique est en général convoitée par deux types de tâches, l'une tend à faire exercer l'élève le passage du registre langagier au symbolique et vice-versa, l'autre vise faire pratiquer les axiomes de Kolmogorov.

Encore une fois, la faible fréquentation de plusieurs modalités originellement dénommées *Dessins, Graphiques* et *Arbres* nous a obligé à les regrouper dans une nouvelle modalité dénommée *Autre*. Cette catégorie réunit des modalités conceptuellement bien différentes par sa fonction, par exemple le registre graphique qui semble favoriser la lecture et la mise en relation appartient à cette nouvelle modalité de même qu'un diagramme comme un arbre qui a tendance à être utilisé comme support au registre numérique. Néanmoins, même si hétérogène, cette modalité reste marginale en termes d'effectifs (Tableau 9)



Graphique 8

Distribution d'exercices par type de représentation (effectifs et pourcentage)

Manuel	Présence objets visuels								Total
	Tableau		Autre		Symbol		Aucun		
BELIN 1S	*18*	*17,3%*	*12*	*11,5%*	*21*	*20,2%*	*53*	*51,0%*	104
NATHAN 1S	*7*	*9,6%*	*4*	*5,5%*	*15*	*20,5%*	*47*	*64,4%*	73
BREAL ES	*6*	*19,4%*	*0*	*0,0%*	*2*	*6,5%*	*23*	*74,2%*	31
DIDIER ES	*29*	*44,6%*	*6*	*9,2%*	*5*	*7,7%*	*25*	*38,5%*	65

Tableau 9

Tâche Principale

Nous avons constaté que les exercices se caractérisent pour être plutôt courts, centrés principalement sur le registre symbolique dont les tableaux servent en général à organiser l'information numérique (probabilités ou cardinaux de sous ensembles). Avec cette nouvelle variable nous nous intéressons à la tâche mathématique principale nécessaire à la résolution d'un exercice. Cette variable sera confrontée à d'autres qui s'adressent à des aspects interprétatifs et fonctionnels en termes de dialectique outil-objet. Par exemple, avec cette variable nous essayons de repérer si un exercice propose à l'élève de déterminer l'espérance mathématique, avec d'autres variables, nous observerons si cet objet calculé est un outil ou non pour une prise de décision.

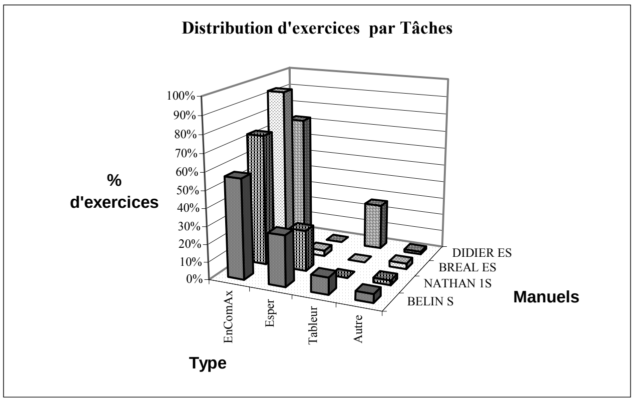
En principe les modalités étaient plus nombreuses, un regroupement a réuni quelques-unes relativement proches (*Ensembliste, Combinatoire* et *Axiomes*). La première représentait des tâches liées à notions de la théorie d'ensembles, la deuxième de combinatoire, etc.

Une tâche, bien que non mathématique *stricto sensu* est appelée *Tableur :* nous y condensons toutes celles qui s'occupent de faire travailler l'élève sur des notions liées à une interface informatique telle que le tableur.

Le Graphique 9 montre une distribution fortement concentrée sur la modalité *EnComAx* (Ensembliste, Combinatoire, Axiomes). Les manuels de la filière S font intervenir des tâches plus complexes que celles de la filière ES, tel le cas l'espérance mathématique (*Esper*). En compensation, ceux de la filière ES choisissent soit d'approfondir le travail sur la théorie d'ensembles et les axiomes de Kolmogorov (Bréal ES, plus de 93 % de ses exercices, Tableau 10) soit de renforcer le travail sur des outils TICE.

Ainsi l'écologie de la probabilité (Chevallard, 1985) semble assez restreinte et polarisée sur des tâches de la première modalité. Une voie alternative à cette polarisation semble être le travail sur des aspects informatiques. Les chapitres de Statistique pourraient donc contribuer à détendre la tension provoquée par la demande d'introduction de nouvelles technologies dans l'enseignement des mathématiques, il reste à considérer le risque de glissement de problématique où un travail statistique pourrait être déplacé pour un de type informatique.

Par l'introduction d'autres variables, nous nous intéresserons à observer si le manque de diversité de travail mathématique est justifié par un travail statistique interprétatif ou décisionnel.



Graphique 9

Distribution d'exercices par tâche (effectifs et pourcentage)

Manuel	Tâche: Ens-Comb-Axiom		Esper		Tableur		Autre		Total
BELIN 1S	*59*	*56,7%*	*30*	*28,8%*	*10*	*9,6%*	*5*	*4,8%*	104
NATHAN 1S	*54*	*74,0%*	*17*	*23,3%*	*0*	*0,0%*	*2*	*2,7%*	73
BREAL ES	*29*	*93,5%*	*1*	*3,2%*	*0*	*0,0%*	*1*	*3,2%*	31
DIDIER ES	*47*	*72,3%*	*0*	*0,0%*	*17*	*26,2%*	*1*	*1,5%*	65

Tableau 10

Variables concernant la probabilité

Type de Contexte

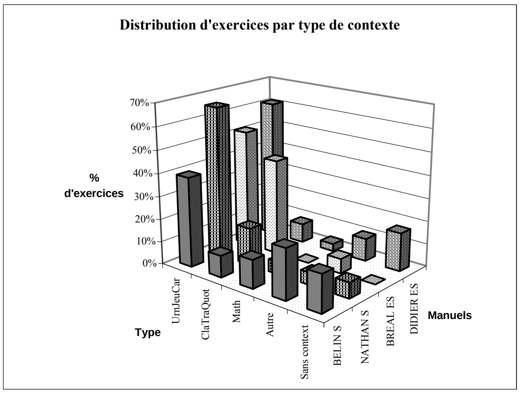
Nous nous interrogeons aussi sur le type de contextes proposés dans les exercices du chapitre Probabilité. Si un profil était dévoilé, la probabilité serait ainsi implicitement associée à ce genre de contextes, en termes probablement de la théorie des champs conceptuels ((Vergnaud, 1990)), les élèves retiendraient lors de leur passage pour la classe de Première, que la probabilité donnerait des réponses à un certain genre de problèmes, ceux du profil en question.

Bien que la probabilité soit un outil de vaste application, nous observons dans ces manuels qu'elle reste restreinte en diversité. Les modalités se résument principalement à trois catégories (*Urne-jeux-carte*, *Classe-Travail-Quotidien, Autre),* les deux dernières étant le fruit d'une réorganisation. La modalité *Classe-Travail-Quotidien* concentre tous les contextes (toujours évoqués) d'une réalité scolaire ou liée au monde du travail ou à des questions de la vie quotidienne. La modalité *Mathématiques* réunit des exercices dont le contexte est le plan abstrait mathématique, comme par exemple, estimer l'aire sous une courbe parabolique par la proportion de points qui tombent sous la courbe lorsqu 'ils sont lancés par simulation de manière uniforme sur le plan.

Le Graphique 10 met en évidence un choix assez prononcé des rédacteurs de ces manuels : la modalité *Urne-jeux-cartes* est d'une part la plus précise dans le niveau hiérarchique conceptuel (elle n'est pas le fruit d'un regroupement) et d'autre part, la plus fréquente dans les manuels, et cela sans distinction de filière.

En principe, les objets des contextes de cette modalité pourraient se présenter de deux manières, l'une directe, en tant que contexte proprement dit de l'exercice, l'autre indirecte, comme une pseudo-réalité intermédiaire d'une autre plus complexe. Dans tous les exercices de ces quatre manuels, ces objets participent de manière directe, en tant que réalité en soi et non pas comme une pseudo-réalité modélisante.

D'ailleurs, la forte présence de contextes ludiques nous interroge sur le type d'utilités que pourraient retenir les élèves de la notion de probabilité; un objet que d'après les manuels semblerait servir principalement pour agir en situations de divertissement ou passe-temps. Un outil applicable à des activités oisives dont leur perception sociale semble être, selon les sociétés, plus ou moins négative.



Graphique 10

Distribution d'exercices par type de contexte (effectifs et pourcentage)

Manuel	Type contexte										Total
	urnes-jeux-cartes		Clas-Trav-Quot		Math		Autre		Sans context		
BELIN 1S	*41*	*39,4%*	*10*	*9,6%*	*13*	*12,5%*	*23*	*22,1%*	*17*	*16,3%*	104
NATHAN 1S	*48*	*65,8%*	*12*	*16,4%*	*4*	*5,5%*	*4*	*5,5%*	*5*	*6,8%*	73
BREAL ES	*16*	*51,6%*	*13*	*41,9%*	*0*	*0,0%*	*2*	*6,5%*	*0*	*0,0%*	31
DIDIER ES	*36*	*61,0%*	*5*	*8,5%*	*2*	*3,4%*	*6*	*10,2%*	*10*	*16,9%*	59

Tableau 11

Dan le premier chapitre nous avons proposé un ensemble d'éléments caractéristiques pour chaque interprétation de la probabilité. Cette recherche d'éléments caractéristiques était motivée, entre autres, par l'intention de montrer que l'interprétation d'une probabilité n'est pas un choix subjectif, mais une propriété du couple contexte/question posée, cette sorte d'objectivité fait que l'interprétation sous-jacente soit un élément observable dans l'exercice. Sans une correspondance entre l'interprétation de la probabilité et les éléments caractéristiques, il nous serait impossible de préciser l'interprétation sous-jacente à un exercice quelconque. Les catégories proposées dans notre caractérisation tendent précisément à fournir des éléments objectifs permettant l'identification de l'interprétation de la probabilité à un problème donné.

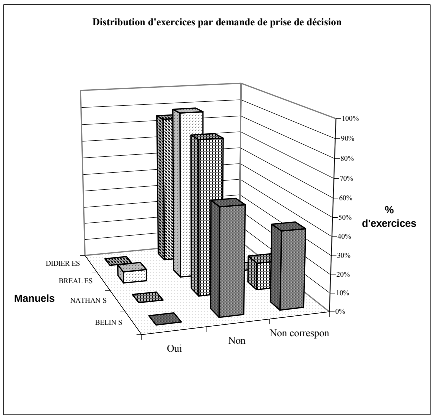
D'ailleurs, si un problème manque de contexte, il deviendrait, selon notre approche, impossible de lui associer une interprétation à la probabilité, tel est le cas des exercices de la modalité *Sans contexte* dont effectivement la probabilité opère dans l'abstraction. Cette modalité n'est pas fréquentée de la même manière pour ces quatre manuels, même si elle n'est exclusive à aucune filière en particulier.

Bien qu'il soit possible d'attacher avec une relative précision une interprétation à un exercice donné, cela n'implique pas que l'interprétation soit nécessaire à sa résolution. Pour cela la question posée doit évidement concerner une approche interprétative et non exclusivement calculatoire. Nous avons introduit une variable qui tente de repérer cette problématique, nous reviendrons plus tard sur cette question.

Décision

La probabilité peut être considérée comme un outil de prise de décision, un des plus élémentaires, certes. Les quelques exemples cités dans le premier chapitre témoignent de la diversité de son champ d'application. La probabilité est en effet un élément clé pour décider rationnellement dans un contexte d'incertitude, même certains épistémologues lui attribuent le début d'une rationalité sur l'incertain (Hacking, 2002).

Nous nous interrogeons donc sur la place de la probabilité comme outil décisionnel dans les exercices de ces manuels. En d'autres termes, nous souhaitons savoir s'il y a une demande de prise de décision à partir de la probabilité calculée. Nous avons retenus trois modalités pour cette question (*Oui, Non* et *Non correspond*). La troisième concerne, entre autres, aux exercices qui manquent de contexte. En effet, sans contexte il n'y aurait aucune décision à prendre.



Graphique 11

Les deux problèmes du manuel Bréal S (Exercices 11 et 14) que nous avons associé à la modalité *Oui* ne correspondent pas strictement à une prise de décision à partir de la probabilité calculée. Malgré cela, nous avons conservé ces deux observations et nous les avons placées dans cette modalité pour éviter de créer des modalités à faibles effectifs. Ces deux exercices mobilisent les principes d'une inférence sur la proportion de la population par la vraisemblance de la proportion de l'échantillon.

Le manuel Belin de la filière S se différencie par le nombre d'exercices sans contexte (44/104 exercices), la tâche consistant notamment à calculer la probabilité par l'application d'axiomes.

La tendance, les quatre manuels confondus, n'est pas de traiter la probabilité comme un outil de prise de décision. Les exercices se caractérisent plutôt pour la considérer comme un objet de calcul. Nous verrons que la demande de son évaluation numérique est presque toujours explicite, finissant l'exercice lors de son calcul (ou du calcul de l'espérance mathématique), mais en général la probabilité ne dépasse l'intérêt calculatoire. Il s'agit de préciser les cas élémentaires, puis les possibles pour enfin établir le quotient des deux ensembles. Les principales difficultés résidant dans la détermination de leurs cardinaux.

Distribution d'exercices par demande de prise de décision (effectifs et pourcentage)

Manuel	Décision						Total
	Oui		Non		Non correspond		
BELIN S	*0*	*0,0 %*	*60*	*57,7%*	*44*	*42,3%*	104
NATHAN S	*0*	*0,0%*	*62*	*84,9%*	*11*	*15,1%*	73
BREAL ES	*2*	*6,5%*	*29*	*93,5%*	*0*	*0,0%*	31
DIDIER ES	*0*	*0,0%*	*55*	*84,6%*	*10*	*15,4%*	65

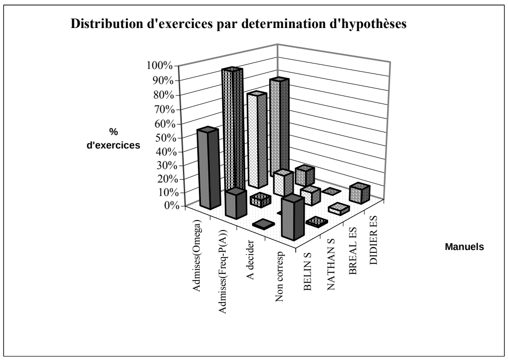
Tableau 12

Hypothèses du modèle

Cette autre variable d'une part fonctionne comme élément de contrôle de la cohérence d'autres variables observées et d'autre part cherche à décrire les moyens mis à disposition pour évaluer numériquement une probabilité. Rappelons que l'évaluation numérique *a priori* de la notion fréquentiste requiert l'assomption d'un ensemble d'hypothèses, et que l'évaluation de la bayésienne peut s'effectuer par des moyens plus divers (entre autres, principe fréquentiste, d'équipossibilité). Les modalités pour cette variable sont :

- *Admises(Omega).* Lorsque l'exercice fourni un ensemble fondamental de probabilité permettant de déterminer les cardinaux des ensembles intervenant dans le calcul de la probabilité. Ces exercices se basent sur le principe d'équipossibilité de cas élémentaires.
- *Admises (Fréquence ou P(A)).* Lorsque l'exercice précise dans son énoncé la valeur de la probabilité par l'estimation d'un échantillon où directement en la donnant sans préciser son origine.
- *A décider.* Lorsque l'exercice demande d'estimer la valeur d'un paramètre de la population. Cette modalité impliquerait un test d'hypothèse (bayésien ou fréquentiste), ces concepts sont développés en classe de Terminal, nous n'attendons pas à observer des exercices qui demandent une telle formalité en classe de Première.
- *Non correspond*. Ces sont les exercices pour lesquels une telle question ne correspond pas, par exemple si un exercice ne présente pas un contexte il n'y aurait pas non plus d'hypothèses.

Cette variable sert d'une part donc à confirmer d'autres en agissant comme une corroboration des modalités observées, d'autre part elle nous renseigne sur la manière que les manuels proposent aux élèves de déterminer les probabilités soit pour une épreuve générique (probabilité bayésienne) soit pour la série infinie (probabilité fréquentiste). Le cas de probabiliser sur une hypothèse, correspond à la notion bayésienne et est généralement appliqué par la formule de Bayes, notion prévue pour la classe de Terminale.



Graphique 12

La variable *Type de contexte* (Graphique 10) nous informait de la forte concentration d'exercices sur la modalité jeux de cartes, urnes et dés. Ces exercices semblent faciliter de manière presque immédiate l'admission d'hypothèses d'équipossibilité de cas élémentaires, caractéristique mise en évidence par la modalité *Hypothèses Admises(Omega)* (Graphique 12). Cette modalité nous renseigne de la forte mobilisation (entre 55,8% et 93,2% des exercices par manuel, Tableau 13) du principe d'équipossibilité et de son application en situations assez similaires (situations ludiques) et cela sans distinction de filière. La correspondance entre ces modalités des variables *Type de contexte* et *Hypothèse du modèle* est déductible, d'autres relations quasi implicatives seront traitées avec le logiciel CHIC pour étudier les flux entre modalités de variables différentes.

D'autres variables nous aideront à identifier sur quel type de probabilité s'appliqueront ces ensembles de cas élémentaires équiprobables. En fait, la modalité *Hypothèses admises(Omega)* ne veut pas dire que la probabilité sera du type fréquentiste, pour cela il faudra considérer l'objet sur lequel porte la probabilité, si c'est une épreuve générique, nous serons devant une probabilité bayésienne dont la proportion de cas favorables sur cas possibles est une raison de croire basée sur le principe d'équipossibilité ; si l'objet est la série infinie il s'agirait effectivement d'une probabilité fréquentiste.

Distribution d'exercices par demande de prise de décision (effectifs et pourcentage)

Manuel	Hypothèses modèle								Total
	Admises (Omega)		Admises (Freq-P(A))		A décider		Non corresp		
BELIN S	*58*	*55,8%*	*18*	*17,3%*	*1*	*1,0%*	*27*	*26,0%*	104
NATHAN S	*68*	*93,2%*	*4*	*5,5%*	*0*	*0,0%*	*1*	*1,4%*	73
BREAL ES	*22*	*71,0%*	*5*	*16,1%*	*3*	*9,7%*	*1*	*3,2%*	31
DIDIER ES	*50*	*76,9%*	*8*	*12,3%*	*0*	*0,0%*	*7*	*10,8%*	65

Tableau 13

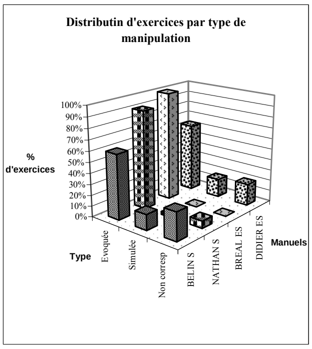
Manipulation d'objets

Cette variable décline en trois modalités : *Manipulation Evoquée*, S*imulée* et *Non correspond,* soulignons qu'une modalité appelée *Réelle* avait été considérée et puis écartée par la nullité des effectifs pour les quatre manuels. Cette modalité représentait les exercices qui proposaient à l'élève d'effectuer une expérimentation concrète.

Nous avons introduit cette variable pour trois raisons :

- La première est en rapport au rôle des simulations dans la probabilité fréquentiste. Une simulation est une représentation artificielle d'un phénomène réel. Si le phénomène auquel la simulation fait référence n'a pas été précédemment appris et les liens entre lui et son évocation n'ont pas été construits, la simulation risquerait de perdre sa place de représentation en se constituant en réalité en soi. De cette manière les conclusions qui y seraient tirées ne seraient plus transposables à la réalité de référence. Avec cette variable nous tentons de repérer sur lequel de ces deux pôles (réalité – simulation) se place un exercice.
- La deuxième concerne la mobilisation de représentations déterministes. La réalité est la principale source d'indéterminisme, et nous ne sommes pas sûrs que les évocations ou les simulations soient des représentations équivalentes à la réalité pour une remise en question de représentations déterministes.
- La troisième se réfère à l'approche bayésienne. Vue que la probabilité est un degré de certitude, la qualité de l'information semblerait jouer un rôle dans l'évaluation personnelle de la probabilité (Gärdenfors et al., 1988), pages 12, 28, 94 et 98 ; (Hacking, 2002), pages 133 et 188). Par exemple, dans une probabilité conditionnelle ($P_B(A)$, dont B est une information conditionnant le degré de certitude sur A) l'évocation de B pourrait influer différemment sur l'évaluation de la probabilité d'A que le ferait sa constatation personnelle. Nous essayons d'observer si les informations varient donc en qualité ou non.

Le Graphique 13 montre la distribution des exercices des quatre manuels où d'une part on constate une homogénéité par des situations évoquées et d'autre part un faible taux de simulations. D'ailleurs, aucune différenciation n'est envisageable entre les filières : les manuels Nathan S et Bréal ES se rapprochent entre eux du même Belin S et Didier ES en fonction des trois modalités considérées.



Graphique 13

La modalité écartée (*Manipulation réelle*) attire l'attention par son absence, aucun manuel n'a proposé des exercices demandant à l'élève d'effectuer des expérimentations. Les questions qui soulèvent les trois raisons détaillées ci-dessus restent une question ouverte dans cette analyse d'espace potentiel de travail du chapitre Probabilité de ces quatre manuels. Premièrement, aucun manuel n'a proposé de réunir expérimentation et simulation dans le même exercice, le lien entre situation évoquant et évoquée qu'à notre avis pourrait ne pas être transparent, n'a pas été renforcé dans ces quatre manuels, deuxièmement la qualité de l'information intervenant dans ces exercices nous interpelle, les données n'ayant jamais été produites par l'élève. Troisièmement, due à nous doutes concernant l'équivalence entre situation réelle et sa simulation, et vu le manque d'exercices de contexte réel, nous nous interrogeons sur la manière en que ces exercices remettraient en question les conceptions déterministes des élèves.

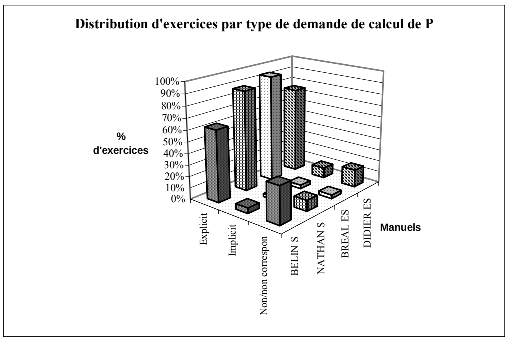
Distribution d'exercices par type de manipulation d'objets (effectifs et pourcentage)

Manuel	Manipulation						Total
	Evoquée		Simulée		Non correspon		
BELIN S	*62*	*59,6%*	*15*	*14,4%*	*27*	*26,0%*	104
NATHAN S	*66*	*90,4%*	*2*	*2,7%*	*5*	*6,8%*	73
BREAL ES	*31*	*100,0%*	*0*	*0,0%*	*0*	*0,0%*	31
DIDIER ES	*41*	*63,1%*	*11*	*16,9%*	*13*	*20,0%*	65

Tableau 14

Calcul de P

Avec cette variable nous essayons de complémenter l'information fournie par d'autres. Nous avons observé que les exercices ne se servent pas de la probabilité pour une décision et qu'en général ils concernent des situations ludiques évoquées. La variable *Calcul de P* cherche à observer la manière dont l'élève est invité à évaluer une probabilité, en d'autres termes, si son calcul est à mobiliser (modalité *Demande implicite*) ou si par contre l'élève arrive à cette tâche avec l'intention de répondre à une *Demande explicite* de l'exercice. Une troisième modalité est considérée pour les exercices dont soit le calcul n'est pas demandé soit les caractéristiques de l'exercice ne permettent pas ce passage au registre numérique (modalité *Non-Non correspond*).



Graphique 14

Comme d'autres, la variable *Calcul de P* ne permet pas une différenciation entre filières (Graphique 14), bien au contraire elle réunit par paires les manuels de filières différentes (Nathan S avec Bréal ES et Belin S avec Didier ES). De leur part, es manuels Belin S et Didier ES, choisissent de proposer des exercices sans demande de calcul de probabilité, qui sont en général des exercices d'entraînement.

Les exercices, faibles en nombre, qui recourent à une demande implicite du calcul de la probabilité le font :

- Par la demande, cette fois-ci explicite, du calcul d'un autre objet (l'espérance mathématique). Principalement chez les manuels de la filière S (en particulier Belin S) (Tableau 15 et Graphique 9). Ces calculs n'arrivent pas à être soumis à des interprétations.
- D'une manière plus diverse pour l'associer à la stabilisation de fréquence d'une simulation par l'emploi de termes synonymiques tels que chance ou rareté d'un phénomène (principalement Didier ES) (Graphique 9 et Graphique 13). Ces exercices sont les moins nombreux.

La séquence caractéristique des exercices de la modalité *Demande explicite* peut être résumée de la manière suivante :

- Description d'un contexte de tirage au sort (jeux, urnes, dés).
- Description de l'évènement visant sa cardinalité (Hypothèses admises par Ensembles finis)
- Demande explicite du calcul de la probabilité
- Fin de l'exercice

Dans cette séquence, les difficultés se concentrent généralement sur la détermination des cardinaux des ensembles, ou sur les axiomes de la théorie de Kolmogorov.

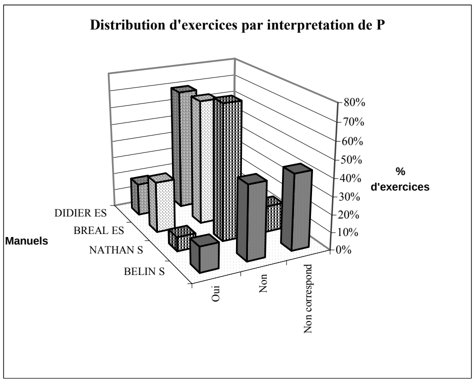
Distribution d'exercices par type de demande de calcul de P (effectifs et pourcentage)

Manuel	Calcul P						
	Explicit		Implicite		Non/non correspond		Total
BELIN S	*65*	*62,5%*	*5*	*4,8%*	*34*	*32,7%*	104
NATHAN S	*64*	*87,7%*	*2*	*2,7%*	*7*	*9,6%*	73
BREAL ES	*29*	*93,5%*	*1*	*3,2%*	*1*	*3,2%*	31
DIDIER ES	*49*	*75,4%*	*6*	*9,2%*	*10*	*15,4%*	65

Tableau 15

Interprétation de P

La variable précédente sert à nous informer sur le passage de la probabilité par le registre numérique. Maintenant nous nous intéressons à son possible retour au registre langagier lors de son calcul. Les trois modalités retenues sont : *Oui, Non* et *Non correspond.* En principe la modalité *Oui* déclinait en *Explicitement* et *Implicitement,* mais encore une fois leur faible nombre d'effectifs nous a mené à les à regrouper.



Graphique 15

Le Graphique 15 montre encore une fois une forte concentration sur une modalité, cette fois-ci sur celle qui caractérise les exercices par la non interprétation de la probabilité calculée, les pourcentages varient entre 42,3% (Belin S, 44/104 exercices) et 76,7% (Nathan S, 56/73), le taux relativement faible chez Belin S est dû au choix de ses auteurs de proposer un nombre important d'exercices sans contextes et d'entraînement (45/104 exercices, Tableau 16). Si nous réunissons les exercices qui pour une raison ou une autre ne font pas intervenir une interprétation de probabilité, les pourcentages augmentent fortement, variant entre 71 % (Bréal ES, 22/31 exercices) et 91,8% (Nathan S, 67/73 exercices).

Distribution d'exercices par interprétation de P (effectifs et pourcentage)

Manuel	Interprétation P						
	Oui		Non		Non correspond		Total
BELIN S	*15*	*14,4%*	*44*	*42,3%*	*45*	*43,3%*	104
NATHAN S	*6*	*8,2%*	*56*	*76,7%*	*11*	*15,1%*	73
BREAL ES	*9*	*29,0%*	*22*	*71,0%*	*0*	*0,0%*	31
DIDIER ES	*12*	*18,5%*	*45*	*69,2%*	*8*	*12,3%*	65

Tableau 16

Nous introduirons une dernière variable qui nous renseignera la nature de l'objet sur lequel porte la probabilité, ceci nous permettra de déterminer le type d'interprétation sous-jacente au problème en fonction des caractéristiques de son contexte.

Nature d'A

Nous venons d'observer d'une part, une forte proportion d'exercices demandant le calcul d'une probabilité, et d'autre part un faible pourcentage d'exercices associant ce calcul à une interprétation. Ces pourcentages varient entre 8,2 % (Nathan S, 6/73 exercices) et 29% (Bréal ES, 9/31 exercices). Avec cette nouvelle variable dénommée *Nature d'A,* nous cherchons à examiner la nature de l'objet sur lequel s'effectue le calcul de la probabilité et déterminer ainsi la notion de la probabilité sous-jacente à l'exercice. Selon la *Nature d'A* nous déterminerons le type d'interprétation associée à un exercice.

Rappelons que nos éléments caractéristiques visent à fournir des critères objectifs permettant de déterminer l'interprétation associée à un contexte en fonction de la question posée, ces éléments caractéristiques se complémentent entre eux cherchant à créer un ensemble cohérent d'analyse. Par exemple, un élément caractéristique de la probabilité fréquentiste est le raisonnement déductif, ce genre de démarche argumentative évalue la fausseté ou véracité d'une proposition (valeurs logiques associés à la notion fréquentiste de la probabilité).

Il serait donc contradictoire d'affirmer que la probabilité fréquentiste mobilise un raisonnement déductif et puis déclarer qu'une épreuve générique est un cas fréquentiste. La valeur logique d'une épreuve générique n'est pas du type dichotomique (vrai faux) mais dans un intervalle continu entre le vrai et le faux, ces valeurs logiques ne correspondant pas avec le raisonnement déductif.

Voilà donc le genre de cohérence entre les raisonnements et les valeurs logiques associées. Ces idées ont été discutées dans notre premier chapitre, nous nous servirons en particulier de l'élément *Nature d'A*.

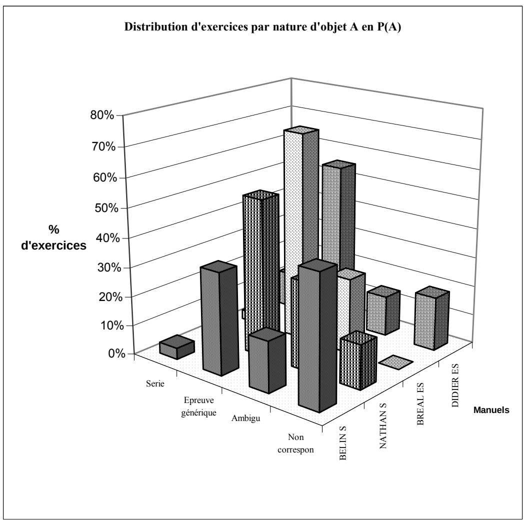
Les modalités pour la variable *Nature de l'objet à probabiliser* étaient en principe cinq : *Hypothèses, Epreuve générique, Série, Ambiguë* et *Non correspond.* Dans le Chapitre I, nous avons décrit un critère d'évaluation permettant le passage au nombre d'une probabilité portée sur une hypothèse, c'était le principe de raison insuffisante. Nous n'avons repéré aucun exercice le faisant intervenir, ni non plus des exercices qui évaluent des probabilités sur des hypothèses en se servant d'autres critères. Dû à ses effectifs nuls, ce genre de propositions n'a pas été retenu dans notre analyse. Finalement, les modalités sont :

- *Série.* Une série infinie est caractérisée premièrement par la convergence d'apparition d'un phénomène lors de la répétition de l'épreuve sous les mêmes conditions jusqu'à l'infini et deuxième par la randomisation de la série (Chapitre I. Richard Von Mises. La probabilité fréquentiste).
- *Epreuve générique.* Les épreuves génériques sont caractérisées par sa plausibilité de reproduction sous des conditions considérées comme étant identiques. La principale différence entre une épreuve générique et la série infinie réside dans le fait que la première s'intéresse à une seule réalisation, la deuxième à un nombre infini.
- *Non corresponde.* Un exercice appartient à cette modalité s'il n'y a pas de contexte ou il n'y a pas de calcul de probabilité concerné.
- *Ambiguë.* A cette modalité floue appartiennent les exercices pour lesquels la nature de l'objet en question n'est pas claire. Cette modalité est fréquentée, entre autres, par des exercices du type épreuve générique dont l'évaluation numérique se fait par le principe fréquentiste mais le rôle de cette dernière en tant que raison de croire est surévaluée et devient interprétation en soi. Nous avons décrit ce procédé dans le Chapitre I, page 83 (*Conclusions. Des éléments communs dans leurs manières de les évaluer*, voir aussi Figures 5, 6 et 7). A cette modalité appartiennent aussi les exercices qui demandent de déterminer "une loi de probabilité", cette expression en termes abstraits ne permet pas de déduire la nature de l'objet sur lequel s'évalue la probabilité. Enfin, à elle appartiennent donc les exercices dont il n'y a pas d'indices précis de la nature de l'objet sur lequel porte le calcul de la probabilité.

Le Graphique 16 représente la distribution en pourcentage des exercices par nature de l'objet et manuel. Les exercices de la modalité *Série*, autrement dit ceux qui correspondent à une interprétation fréquentiste sont les moins nombreux, les pourcentages varient entre 1,4 % (Nathan S, 1/73 exercices) et 12,3 % (Didier ES, 8/65 exercices), ces exercices, rares, seraient par leurs contextes, cohérents avec l'interprétation fréquentiste. Ils seraient les seuls qui correspondent avec l'interprétation admise par les programmes officiels. L'analyse implicative nous permettra de déterminer si pour ces exercices de nature fréquentiste une interprétation a été demandée ou non.

La modalité *Ambiguë* avec des pourcentages entre 13, 8 % (Didier ES, 9/65 exercices) et 30,1 % (Nathan S, 22/73 exercices) réunit les exercices du type épreuve générique où l'évaluation numérique se fait par le principe fréquentiste, le vrai rôle de la fréquence d'apparition de l'événement restant caché. En d'autres termes, un exercice de cette modalité demande en général de calculer une probabilité sur des hypothèses d'équipossibilité et puis de valider ou réfuter cette valeur par une simulation sur un logiciel tableur. Dans ce genre de cas, les deux interprétations se sont vues inversées leurs rôles, celle qui est la notion centrale reste "cachée", celle qui opère comme une raison de croire est mise au premier plan devenant le cas échéant l'objet à institutionnaliser.

Les *Epreuves génériques* occupent une place centrale dans les exercices de ces quatre manuels, avec des pourcentages entre 34,6% (Belin S, 37/104 exercices) et 71 % (Bréal ES, 22/31 exercices). Les contextes de ces problèmes répondent à l'approche bayésienne dont leur évaluation se fait par un ensemble fini de référence. Comme ces exercices sont en général du type jeux de cartes, urnes ou dés (nous le confirmeront par l'analyse implicatif), sa reproductibilité est non seulement plausible mais de conception (mentale, tous les exercices sont évocateurs des contextes) immédiate. Cette possibilité de reproduction immédiate permet aisément d'appliquer le principe fréquentiste. Ainsi lorsqu'une interprétation doit s'expliciter pour ces exercices il existe toujours la possibilité, par un arrangement, de leur attacher la fréquence d'apparition, même si elle opère de manière secondaire. A la différence de ceux de la modalité *Ambiguë*, ces exercices ne font aucune mention de l'utilisation du principe fréquentiste, ils mobilisent tous le principe d'équipossibilité.



Graphique 16

Les *épreuves génériques* constituent le contexte le plus fréquenté, et cela pour les quatre manuels. L'observation de tant de contextes de nature bayésienne devrait attirer l'attention sur le paradoxe qui représente le constat semblant contraire aux directives des documents officiels : les élèves sont confrontés plus fréquemment à des situations bayésiennes que fréquentistes. Nous nous servirons de l'analyse conjointe de ces variables pour comprendre ce phénomène si particulier dans l'enseignement de la probabilité.

Toutefois, quelques explications à ce paradoxe se trouvent dans les caractéristiques mêmes des épreuves génériques. Les analyses *a priori* de ces exercices nous suggèrent des contextes similaires en structure, simples à énoncer et précises lors de la définition des objets nécessaires au calcul d'une probabilité. Ces caractéristiques pourraient devenir des avantages pour la gestion de la classe, des caractéristiques qui se traduiraient en classe en tâches courtes en termes de temps de travail, d'institutionnalisations ponctuelles et concises. L'exemple suivant illustre ces propriétés didactiques de ce genre d'exercices (Belin S, page 211, exercice 60) :

> *«On jette (...) trois fois un dé équilibré dont les six faces sont numérotées de 1 à 6 et on note les résultats obtenus (...) Calculer les probabilités des événements :*
> *A : "les trois nombres tirés sont pairs"   (...) »*

Nous avons en quelques lignes un exercice complet. Il permet, dans un espace réduit, de définir tous les éléments nécessaires aux calculs. Même plus, en jouant sur la définition de l'événement on peut faire travailler les élèves sur un nombre important de stratégies différentes de calculs, même combinatoires. Les épreuves génériques permettent donc de développer un nombre non négligeable de tâches mathématiques calculatoires et algébriques à un coût de gestion réduit et ce n'est pas leur seule vertu, par leur latente possibilité de reproduction elles permettent, si une interprétation s'impose, de recourir à la notion fréquentiste. Ces épreuves génériques seraient donc très profitables, et cela pourrait contribuer à expliquer leur forte présence dans les manuels analysés.

Distribution d'exercices par nature d'objet à probabiliser (effectifs et pourcentage)

Manuel	**Nature de A**								Total
	Série		Epreuve générique		Ambiguë		Non correspond		
BELIN S	4	3,8%	36	34,6%	18	17,3%	46	44,2%	104
NATHAN S	1	1,4%	39	53,4%	22	30,1%	11	15,1%	73
BREAL ES	1	3,2%	22	71,0%	8	25,8%	0	0,0%	31
DIDIER ES	8	12,3%	36	55,4%	9	13,8%	12	18,5%	65

Tableau 17

D'ailleurs, dans un premier moment nous avions prévus plus de variables que celles décrites précédemment, des variables que pour des raisons diverses, nous avons finalement écartés. Une de ces variables était la temporalité de l'objet (ou épreuve) sur lequel porte la probabilité, elle déclinait en plusieurs modalités : passé, présent, et future.

Par exemple, l'exercice cité ci-dessus (Belin S, exercice 60) décrit une situation hypothétiquement déjà accomplie (*les trois nombres tirés sont pairs)*. Si nous admettons le déroulement décrit dans l'énoncé nous devrions nous imaginer l'action des trois lancers déjà accomplisse au moment de nous interroger sur les résultats. Si ces lancers ont été déjà effectués, la situation serait comparable à celle de probabiliser sur une hypothèse. En d'autres termes, nous serions en train de probabiliser sur quelque chose d'achevé, qui ne changera pas, quelque chose que nous est inconnu mais qui a déjà pris sa valeur.

Nous savons qu'il est prohibitif d'un point de vue fréquentiste de probabiliser sur un paramètre d'une population, l'argument pour une telle interdiction se basant sur le fait qu'une hypothèse est fixe et que l'ignorance de sa valeur n'est dûe à un phénomène aléatoire mais épistémique. Il est donc, pour ce paradigme, impossible de probabiliser sur cette valeur.

Alors, si nous admettons que l'épreuve est déjà achevée, et nous nous intéressons à la valeur de cette épreuve, nous pouvons conclure qu'une telle épreuve n'admet être traitée que par l'approche bayésienne.

Dans nos analyses nous nous sommes rendus compte que la variable *Temporalité* et en particulier sa modalité *Passé* a une utilité didactique plutôt que statistique. En d'autres termes, la temporalité contribue à comprendre pourquoi l'épreuve est de nature bayésienne, mais dans aucun cas elle nous a apportée d'information concernant l'exercice en particulier. Ainsi, vu que l'information n'était pas significative, nous l'avons écartée de nos analyses.

La description des variables retenues a été accompagnée d'une lecture de leurs distributions. Pour les quelques conclusions tirées nous avons considéré les variables de manière isolée, dans la deuxième partie de ce chapitre notre analyse sera multidimensionnelle, elle nous permettra d'analyser la relation entre deux variables ou plus, sur la forme de règles implicatives et métarègles et même par la constitution de groupes de variables. Nous commencerons par décrire les principes de l'analyse implicative pour ensuite présenter les interprétations et finalement les conclusions de l'étude des quatre manuels.

Analyse Statistique Implicative

L'analyse statistique implicative est une approche tendant à modéliser l'extraction et la représentation de règles d'inférence partielles entre variables décrivant une population d'individus. Il s'agit pour cette théorie de découvrir des règles non symétriques pour modéliser des relations du type *« si a alors presque b ».*

En autres termes, une population E d'objets est croisée avec des variables que l'on interroge de la façon suivante : « dans quelle mesure peut-on considérer que le fait de relever la variable a impliqué celui de relever de la variable b ? Autrement dit, les sujets ont-ils tendance à être b si l'on sait qu'ils sont a ?

Pour mesurer l'intensité de la quasi implication, cette méthode se base sur une mesure de similarité de I.C. Lerman (Lerman, 1981) pour définir (Gras et al., 1996) une nouvelle mesure, cette fois-ci d'une relation implicative a $\Rightarrow$ b à partir de l'invraisemblance de l'apparition, dans les données du nombre de cas qui l'infirment, c'est-à-dire pour lesquels a est vérifié sans que b ne le soit pas. Cette mesure est relativisée au nombre de données vérifiant a et non b. Elle quantifie " l'étonnement" de l'expert devant le nombre invraisemblablement petit de contre-exemples eu égard à une indépendance et aux effectifs en jeu.

Pour mathématiser la quasi-règle, les auteurs de cette méthode considèrent deux parties quelconques X et Y de l'ensemble E, choisies aléatoirement et indépendamment (absence de lien a priori entre ces deux parties) et de mêmes cardinaux respectifs que A et B. Soit $\overline{Y}$ et $\overline{B}$ les compléments respectifs de Y et de B dans E de même cardinal $n_{\overline{b}} = n - n_b$, dans la Figure 13, les parties hachurées correspondent aux contre-exemples de l'implication a $\Rightarrow$ b.

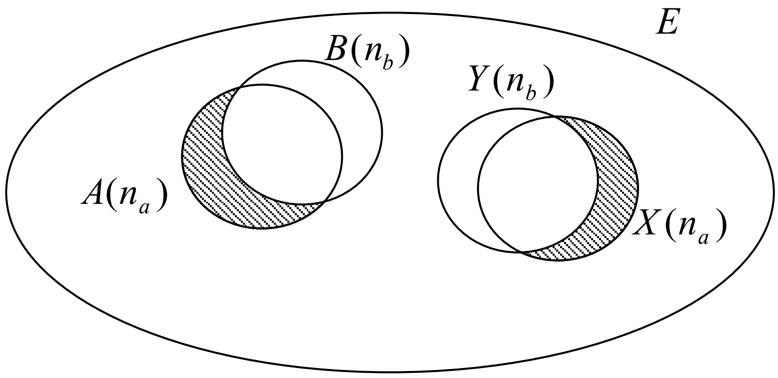


Figure 13

Nous transcrivons un ensemble de définitions de cette théorie (Gras & Bailleul, 2000), page 13):

Définition 1 : a $\Rightarrow$ b est admissible au niveau de confiance 1-α si et seulement si

$$\Pr[\text{card}(X \cap \overline{Y}) \leq card(A \cap \overline{B})] \leq \alpha$$

Il est établit que, pour un certain processus de tirage, la variable aléatoire Card($X \cap \overline{Y}$) suit la loi de Poisson de paramètre $\frac{n_a n_{\overline{b}}}{n}$. Dans les cas où $n_{\overline{b}} \neq 0$ la réduction et la centration de la variable de Poisson devient :

$$Q(a,\overline{b}) = \frac{Card(X \cap \overline{Y}) - \frac{n_a n_{\overline{b}}}{n}}{\sqrt{\frac{n_a n_{\overline{b}}}{n}}}$$

Dans la réalisation expérimentale, la valeur observée de $Q(a,\overline{b})$ est $q(a,\overline{b})$

Définition 2 : $q(a,\overline{b}) = \frac{n_{a \wedge \overline{b}} - \frac{n_a n_{\overline{b}}}{n}}{\sqrt{\frac{n_a n_{\overline{b}}}{n}}}$ est appelé indice d'implication, nombre que nous retenons comme indicateur de la non implication de a sur b.

Définition 3 : Dans le cas où $n_{\overline{b}} \neq$ n l'intensité de l'implication de a sur b est :

$$\varphi(a,b) = 1 - \Pr[Q(a,\overline{b}) \leq q(a,\overline{b})] = \frac{1}{\sqrt{2\pi \int_{q(a,\overline{b})}^{\infty} e^{\frac{t^2}{2}} dt}}$$

Ainsi la définition de l'implication devient :

Définition 4 : L'implication a $\Rightarrow$ b sera admissible au niveau de confiance 1- α si et seulement si $\varphi(a,b) = 1 - \Pr[Q(a,\overline{b}) \leq q(a,\overline{b})] \geq 1 - \alpha$. Cette modélisation de la quasi implication mesure l'étonnement de constater la petitesse des contre-exemples en regard du nombre surprenant des instances de l'implication.

Le logiciel CHIC admet d'autres modélisations que la variable de Poisson, pour notre analyse nous utiliserons la binomiale dû au faible nombre d'effectifs de certaines modalités. Pour cette variable, la modélisation considère les variables duales card($\mathrm{A} \cap \overline{\mathrm{Y}}$) et card($\mathrm{X} \cap \overline{\mathrm{B}}$), où X et Y sont des parties choisies de façon indépendante dans E et respectant les propriétés cardinales respectives de A et B, tout élément de E, par exemple, a la probabilité $\frac{n_a}{n}\frac{n_{\bar{b}}}{n}$ d'appartenir à $A \cap \overline{Y}$. Par suite :

$$\Pr\left[Card(A \cap \overline{Y}) = k\right] = C_n^k \left(\frac{n_a n_{\overline{b}}}{n^2}\right)^k \left(1 - \frac{n_a n_{\overline{b}}}{n^2}\right)^{n-k} = \Pr\left[Card(X \cap \overline{B}) = k\right]$$

Nous avons choisi d'utiliser la méthode classique à la place de l'entropique, vue que cette dernière est plus restrictive et s'avère mieux adaptée à des données plus nombreuses que les nôtres.

Nous utiliserons les trois méthodes disponibles dans le logiciel, ils sont l'analyse de similarité, le graphe implicatif et l'arbre cohésitif. Le premier se base sur une mesure de similarité de I. C. Lerman, les deux autres sur les principes de l'analyse implicative. Pour plus de précision sur ces méthodes nous renvoyons le lecteur aux publications de référence des auteurs de cette théorie (Gras et al., 1996; Gras & Bailleul, 2000).

Nos hypothèses implicatives

Dans le premier chapitre, et à partir d'une enquête épistémologique, nous avons repéré deux grands groupes interprétatifs de la probabilité, l'un bayésien, l'autre fréquentiste. Ces deux approches se sont manifestées tout au long de l'histoire de cette notion. Dans ce deuxième chapitre nous avons montré l'approche des programmes en France, leurs auteurs n'ayant retenu qu'une seule des deux interprétations de la probabilité.

Les textes consultés lors de l'enquête épistémologique nous ont menés à proposer une hypothèse sur la transposition didactique de cette dualité. Cette hypothèse consiste à soutenir que, dû à l'imbrication de signifiés, la dualité de la probabilité deviendrait incontournable lors son enseignement. En d'autres termes, nous admettons que cette dualité de signifiés se reproduirait indéfectiblement dans les situations proposées aux élèves.

L'hypothèse d'incontournable dualité se confirme déjà par la présence non négligeable d'exercices de nature bayésienne, ceci a été observé par la modalité N2 (nature de l'objet épreuve générique) dont sa proportion nous renseigne sur le nombre significatif d'exercices de nature bayésienne dans les quatre manuels analysés.

Néanmoins, cette analyse unidimensionnelle ne nous informe pas des voies entreprises par les manuels pour sortir de cette contradiction. Nous conjecturons que le moyen utilisé pour surmonter cette difficulté est tout simplement un contournement de l'interprétation de la probabilité lorsque elle est de nature bayésienne, en focalisant toute l'attention ses aspects calculatoires.

Cela devra se manifester dans les différentes analyses avec le logiciel CHIC par une absence de lien entre les variables N2 et I1 (objet de nature épreuve générique et interprétation de la probabilité respectivement). Plus précisément, l'analyse de similarités devrait signaler un non rapprochement de ces deux variables, de même que l'implicative ne devrait pas nous renseigner d'implications du type N2 $\Rightarrow$ I1 ou I1 $\Rightarrow$ N2, bien au contraire, la variable N2 devrait être plutôt associée à P1 (Demande explicite de calcul de la probabilité).

Avec cette analyse non seulement nous cherchons à confirmer cette hypothèse d'incontournable dualité, mais elle nous permettra aussi :

- de comprendre comment les manuels résolvent le conflit entre la dualité incontournable et l'approche des programmes.
- d'apprendre sur le type de problèmes auxquels sont confrontés habituellement les élèves, cette connaissance nous permettra de nous approcher des connaissances disponibles chez les élèves de BTS quand nous réaliserons nos expérimentations
- de nous renseigner sur les profils de situations courantes dans le chapitre Probabilités pour en inférer des possibles marges de manœuvre lors de nos expérimentations.

Le logiciel CHIC permet de partitionner les variables en deux groupes, l'un constitué de variables principales, l'autre de secondaires ou supplémentaires. Les premières sont celles sur lesquelles porte l'analyse et qui mettent en évidence les règles et métarègles. Les variables secondaires sont des descripteurs des exercices, elles caractérisent les enchaînements et groupes que les variables principales forment dans les différentes méthodes. Nous présentons le regroupement effectue sur l'ensemble de variables retenues.

Nos variables principales et secondaires

Nos variables supplémentaires ou secondaires constituent, par leur sémantique, des variables extrinsèques ; elles n'interviennent pas directement dans les liaisons exprimées par la classification entre les variables dites principales. Les variables secondaires décrivent passivement les groupes ou classes que les principales définissent activement. Nous avons retenu 20 variables principales et 17 secondaires. Dans nos analyses nous nous servirons indistinctement des termes « modalités » et « variables ». Rappelons que chaque modalité peut être considérée comme une variable binaire.

Variables principales

Décision		
Oui	**Non**	**Non corresp**
D1	**D2**	**D3**

Hypothèses			
Omega	**Fréq-P(A)**	**A décider**	**Non corresp**
H1	**H2**	**H3**	**H4**

Manipulation		
Evoquée	**Simulée**	**Non corresp**
M1	**M2**	**M3**

Calcul de P		
Explicite	**Implicite**	**Non corresp**
P1	**P2**	**P3**

Interprétation de P		
Oui	**Non**	**Non corresp**
I1	**I2**	**I3**

Nature d'A			
Série	**Générique**	**Ambiguë**	**Non corresp**
N1	**N2**	**N3**	**N4**

Variables Secondaires

N° de Lignes	
1-10	**+10**
L1	**L2**

N° de Questions		
1-2	**3-4**	**5-6**
Q1	**Q2**	**Q3**

Type de Représentation			
Tableau	**Autre**	**Symbolique**	**Aucun**
R1	**R2**	**R3**	**R4**

Tâche			
EnsComAx	**Espér**	**Tableur**	**Autre**
T1	**T2**	**T3**	**T4**

Hypothèses			
Omega	**Fréq-P(A)**	**A décider**	**Non corresp**
H1	**H2**	**H3**	**H4**

Contextes				
Jeux	**Trav-Quot**	**Math**	**Autre**	**Sans Cont**
C1	**C2**	**C3**	**C4**	**C5**

Nous effectuons trois analyses par manuel, en commençant par celle de similarités, puis par le graphe implicatif et en finissant par l'arbre cohésitif. Le premier manuel à analyser appartient à la maison Belin et s'adresse aux élèves de la filière Scientifique.

2.5 Manuel Belin S

Ce manuel appartient à la filière Scientifique, pour le chapitre Probabilité il propose 104 exercices repartis en deux sections, la première *Exercices* (Exercices 1 au 92) est suivie par une deuxième appelée *Problèmes* (Exercices 93 au 104). Le Tableau 18 représente la distribution des exercices par modalité. Nous commençons l'analyse statistique par l'arbre de similarités.

Distribution du nombre d'exercices du manuel Belin S par modalité

Lignes		Questions			Représentations				Tâches				Contexte				
L1	L2	Q1	Q2	Q3	R1	R2	R3	R4	T1	T2	T3	T4	C1	C2	C3	C4	C5
75	29	36	26	42	18	12	21	53	59	30	10	5	41	10	13	23	17

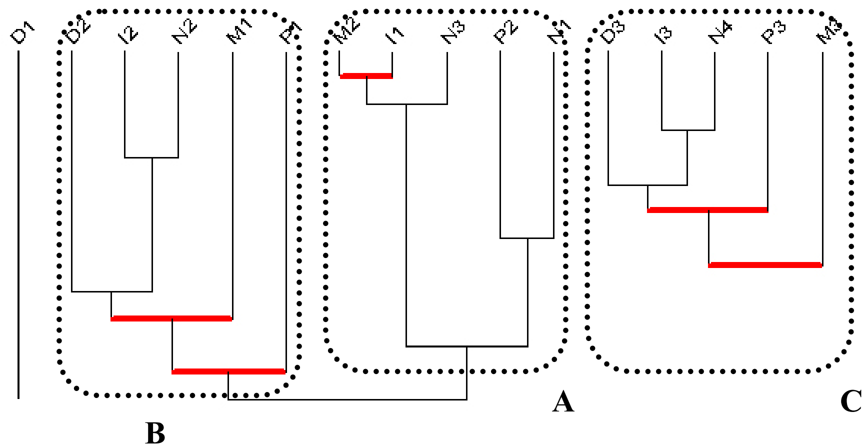
Décision			Hypothèses				Manipulation			Calcul P			Interprétat P			Nature A			
D1	D2	D3	H1	H2	H3	H4	M1	M2	M3	P1	P2	P3	I1	I2	I3	N1	N2	N3	N4
0	60	44	58	18	1	27	62	15	27	65	5	34	15	44	45	4	36	18	46

Tableau 18

Arbre de similarités

Le Graphique 18 met en évidence trois grands groupes que nous les avons appelés A, B et C. Une variable reste isolée (D1) ; elle caractérise les exercices concernant une prise de décision. Les regroupements par variables nous indiquent d'un rassemblement d'exercices selon trois profils différents ; nous analyserons chacun de ces groupes.

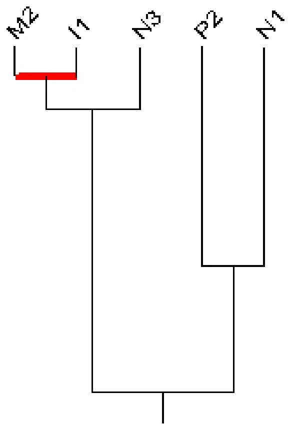
Arbre de similarités. Manuel Belin S



Graphique 18

Groupe A

Le groupe se constitue à partir de quelques modalités associées à la notion fréquentiste. L'ensemble en général nous renseigne sur la manière dont cette notion est traitée chez Belin S. Le Tableau 19 représente le nombre total d'exercices du manuel Belin S concernés pour chacune des modalités retenues, cette classe est d'une part compacte par sa similarité, mais d'autre part faible en nombre d'exercices. En effet, bien qu'elle contienne les modalités les plus proches de l'arbre (M2 et I1, ind. de similarité : 1), le cardinal du premier rassemblement ne représente que 12% des exercices du manuel (groupe optimal classe (M2, I1) : 12 exercices, Card$(M2 \cap I1) = 12$).



Cod	Var	Mod	Card
M2	Manip	Simul	15
I1	Inter P	Oui	15
N3	Nat A	Amb	18
P2	Calc P	Implic	5
N1	Nat A	Série	4

Tableau 19

La modalité N3 rejoint (M2, I1) pour constituer une classe toujours serrée (ind. de similarité 1). Les variables P2 et N1 se rapprochent finalement au groupe (M2,I1) N3). Néanmoins, elles ne constituent pas une sous-classe ni homogène ni significative en nombre (Card$(P2 \cap N1) = 2$), rendant de cette manière la classe hétérogène. En effet :

$$\text{Card}_{M2 \cap I1 \cap N3}(P2 \cap N1) = 0$$

$$\text{Card}_{M2 \cap I1 \cap N3}(\overline{P2} \cap \overline{N1}) = \text{Card}(M2 \cap I1 \cap N3) = 11$$

En fait, les modalités P2 et N1, de si faibles effectifs, rejoignent le groupe A par l'argument inverse : $\text{Card}_{P2 \cap N1}(M2) = \text{Card}(P2 \cap N1) = 2$; $\text{Card}_{P2 \cap N1}(I1) = 1$. La sous-classe précédente ((M2, I1), N3) possède un groupe optimal de 11 exercices, ils se ressemblent par un travail sur la simulation (M2) avec un intérêt pour l'interprétation de la probabilité (I1) dont l'objet sur lequel porte la probabilité est finalement ambigu.

La modalité ambiguë (N3) de ce regroupement nous parle de deux types de situations : soit la description de la nature de l'objet n'est pas précise, soit il y a un concours des deux types d'objets (bayésien et fréquentiste). Dans ce dernier cas, il s'agit en général d'un calcul de probabilité qui porte sur une épreuve générique (bayésienne) évaluée *a priori* par le principe d'équipossibilité et validée *a posteriori* par le principe fréquentiste. Nous reproduirons un exercice extrait du groupe optimal d'individus typiques de la classe ((M2, I1), N3) qui montre le concours des deux interprétations de la probabilité (Belin S, exercices 78, page 213) :

> *On choisi au hasard successivement deux nombres entiers compris entre 0 et 9 inclus. On s'intéresse à l'événement A : « les deux nombres tirés sont pairs ».*
>
> a. *Calculer* $P(A)$ *et* $P(\overline{A})$
>
> b. *On simule 200 fois cette expérience. Déterminer les fréquences des événements* A *et* $\overline{A}$
>
> c. *Comparer les résultats des questions a. et b.*

Notons que dans la première partie, plus précisément jusqu'à l'item a., l'exercice ne propose pas de s'intéresser à la reproduction de l'expérience, même si une telle action est toujours possible. D'après la définition de l'événement A en question on devrait s'intéresser à un phénomène déjà arrivé mais inconnu (« *les deux nombres tirés »)*. L'éventuelle action déjà accomplie nous suggère, par son caractère épistémique, qu'on serait face à une situation bayé sienne, plus précisément il s'agirait d'une épreuve supposée déjà accomplie. Dont, pour l'évaluer on fait recours à un ensemble de référence fini (cas favorables sur possibles). Dans cet exercice l'évaluation numérique de cette épreuve générique se ferait par le principe d'équipossibilité et puis dans les items b. et c. elle serait validée *a posteriori* par la fréquence d'apparition de l'événement. Cet exercice est un exemple de la concurrence de deux notions de la probabilité, des cas que nous avons réunis sous la modalité N3 par l'ambiguïté de la nature de l'objet.

D'ailleurs, cet exercice ne doit pas être considéré comme représentatif de cette classe, au moins en ce qui concerne sa taille. Ils ont tendance à être plutôt longs (L2 t.a.r 0,187) et posant plus de quatre questions (Q3 t.a.r 0.0107). Cet exercice n'est pas représentatif non plus par sa relative richesse de tâches mathématiques concernées (T1 et T2 t.a.r 0,996 et 0,965 respectivement) ; il l'est en tout cas par la mobilisation de ressources informatiques (T3 t.a.r 5.79e-007) et par la mise en disposition dans son énoncé des hypothèses d'équipossibilité (H1 t.a.r 0.0817).

La sous-classe (P2, N1) est la plus faible de l'arbre de classification en termes de cardinal et par son niveau dans la classification par couple de variables (Card($P2 \cap N1) = 2$, niveau 7), elle se caractérise par des exercices dont la nature de l'objet est la série infinie (N1) et le calcul de probabilité est demandé de manière implicite ou indirecte. Il s'agit en général d'exercices longs aussi (L2 t.a.r 0,107), où on demande d'étudier la stabilisation de fréquences par une simulation dans un contexte ludique (C1 t.a.r 0.186), d'hypothèses d'équipossibilités (H1 t.a.r 0,199) ; le calcul de la probabilité est demandé implicitement pour une autre demande, cette fois ci explicite, de calcul de l'espérance et de l'écart type.

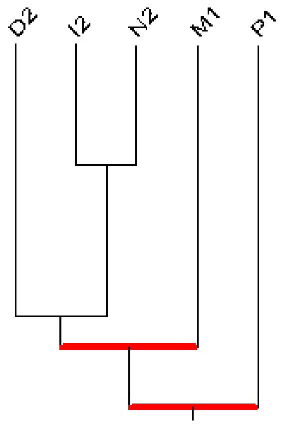
Cette classe qui caractérise les exercices fréquentistes nous renseigne sur une approche faible en nombre d'exercices et en même temps ambiguë. Les exercices de ce groupe ne représentent que 10% des 104 exercices du manuel. La stabilisation de fréquences se manifeste soit comme notion principale (N1), soit ou comme notion secondaire (N3), les premiers cas étant beaucoup plus rares que les deuxièmes où elle apparaît comme raison de croire sous la forme du principe fréquentiste (N3). Ces différents rôles ne sont pas explicités dans les exercices, où il est courant de trouver dans un même énoncé des objets de nature différente.

Finalement, cette notion s'associe à des exercices longs au moins pour deux raisons : d'une part la construction d'un contexte de reproduction sous les mêmes conditions exige l'explicitation d'un nombre non négligeable de conditions. D'autre part, comme toutes ces situations sont simulées, l'exercice doit aussi encadrer le travail sur l'ordinateur.

Groupe B

Le regroupement commence par les variables I2 et N2 (quatrième niveau, ind. de similarité :1). Cette première sous-classe relativement importante en nombre (Card($I2 \cap N2) = 36$) nous renseigne sur les exercices du type épreuve générique (bayésiens) dont il n'est pas nécessaire d'interpréter la probabilité calculée pour le résoudre. A la différence de la précédente, les exercices concernés par cette classe sont plus nombreux (Tableau 20), et en même temps plus proches en similitude (Card($I2 \cap N2 \cap D2 \cap M1 \cap P1$=34, deux nœuds significatifs en niveaux 10 et 12, similarités : 0.997453 et 0.985973 respectivement).

A cette classe se joignent les variables D2, M1 et P1 pour décrire des exercices où l'élève, n'ayant pas une décision à prendre, doit calculer des valeurs de probabilité par une demande explicite, dans les situations toujours évoquées.



Cod	Var	Mod	Card
I2	**Inter P**	**Non**	**44**
N2	**Nat A**	**Génér**	**36**
D2	**Décis**	**Non**	**60**
M1	**Manip**	**Evoq**	**62**
P1	**Calc P**	**Expl**	**65**

Tableau 20

Nous transcrirons à titre d'exemple, un des 36 exercices du groupe optimal de la classe (Exercice 56, Belin S, page 211) :

« On choisi avec équiprobabilité de tirage une carte d'un jeu de 32 cartes.

On considère les événements suivants :

A : « la carte es un cœur »

B : « la carte est un as »

C : « la carte est rouge »

D : « la carte est un valet, une dame ou un roi »

Expliquer chacun des événements suivants et calculer sa probabilité :

$A\cap B$; $A\cap C$; $A\cap D$; $B\cap C$; $B\cap D$; $C\cap D$; $A\cup B$; $A\cup C$; $A\cup D$; $B\cup C$;

$B\cup D$; $C\cup D$ *»*

Encore une fois, l'objet d'intérêt est de nature générique (N2), son caractère facilite sa plausibilité de reproduction mentale. L'élève, d'après le contexte, serait donc interrogé sur la caractéristique de la carta tirée (une carte). La nature de cet objet est renforcée tout au long de l'exercice (« la carte est »…, « la carte est »…).

Dans cet exercice l'interprétation de la probabilité ne joue pas un rôle principal (I2). Qui plus est, elle n'a aucune incidence sur les réponses aux questions posées.

L'incertitude est une condition nécessaire au concept de probabilité. Néanmoins, dans cet exercice l'élève n'est pas invité à agir en se servant de l'outil calculé (D2). Pour résoudre l'exercice, l'élève se représenterait un contexte (ludique d'ailleurs) par son évocation. Il en extrairait des informations pour construire un objet arithmétique qui ne dépassera pas ce statut. En d'autres termes, devant la situation incertaine proposée, l'élève reste spectateur, il contemple, il analyse et il conceptualise des objets qui ne lui serviront pas pour entreprendre une action sur cet environnement.

Les exercices de cette classe ne sont pas toujours longs (L2 t.a.r 0,168), il semblerait que ce soit une caractéristique des épreuves génériques. Ces exercices s'attachent à favoriser les tâches calculatoires et ils sont tendance à être plus courts que les fréquentistes. Pour cette classe ((I2, N2) D2) M1)P1), les exercices posent entre 3 et 4 questions en général (Q2 t.a.r 0,0 343) et le registre principal de représentation reste le langagier (R4 t.a.r 0,00105).

Cette classe se caractérise par le calcul de la probabilité explicitement demandé et qui ne revient pas sur son interprétation. Si la probabilité a un autre intérêt que son propre calcul, cela sera en général parce qu'elle s'insère dans un autre calcul, celui de l'espérance mathématique ou de l'écart type (T2 t.a.r 0,0271). Ces derniers calculs ne pouvant pas être interprétés vu que cela dépend de l'interprétation de la probabilité qu'y intervient.

Ces exercices évoquent des contextes de la vie quotidienne des élèves ou ludiques (C2, t.a.r. 0,00446 ; C1, t.a.r. 0.0211 respectivement) dont les hypothèses *a priori* (H1 t.a.r. 5.55e-006) sont assez évidentes, facilitant ainsi la modélisation du problème par la voie de cas élémentaires équipossibles.

D'ailleurs, bien qu'on puisse considérer chaque exercice comme une unité relativement autonome, nous constatons une relation inter-exercices assez forte. Ce chapitre Probabilité semble organisé par une sorte de groupes hiérarchisés assez bien définis dont une base d'exercices assurerait le traitement de notions nécessaires à d'autres plus complexes jusqu'à un certain niveau où une nouvelle vague d'exercices consolide les bases d'une nouvelle hiérarchie. Ces formations ne sont pas toujours connexes constituant ainsi des blocs d'exercices.

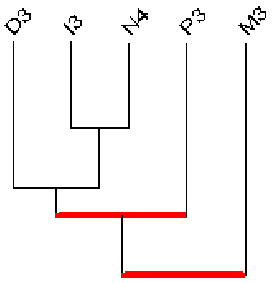
Il est même possible de trouver des traces de cette structure hiérarchique dans le groupe optimal d'individus typiques de cette classe (36 exercices) dont une série d'exercices succesifs (14 au 48) cherche à faire travailler le concept de loi de probabilité, une autre série (56 au 63) les notions ensemblistes, etc.

Les exercices de cette classe ne se focalisent pas sur l'interprétation de la probabilité mais sur des aspects calculatoires mobilisant une relative diversité de notions algébriques ou de combinatoire.

Groupe C

Cette classe, isolée des autres et plus compacte se caractérise par des exercices sans contexte (C5 t.a.r. 1.52e-005). Sans ce dernier, il ne reste que deux possibilités pour les hypothèses, soit elles sont absentes (H4 t.a.r. 5.47e-009) soit elles sont introduites directement sous la forme d'une probabilité P(A) = k (H2 t.a.r. 8.13e-005).

Le manque de contexte fait évidemment qu'il est impossible d'attacher une interprétation au calcul de la probabilité (N3), et cela à cause de la non existence d'objet (N4). L'absence de contexte empêche aussi de s'interroger sur la place de la probabilité comme outil de décision (D3) et finalement d'une éventuelle manipulation (M3).



Cod	**Var**	**Mod**	**Card**
I3	**Inter P**	**Non C**	**45**
N4	**Nat A**	**Non C**	**46**
D3	**Décis**	**Non C**	**44**
P3	**Calc P**	**Non C**	**34**
M3	**Manip**	**Non C**	**27**

Tableau 21

Pour cette classe assez nombreuse (Tableau 21 pour les exercices concernés par variable), il s'agit de faire pratiquer aux élèves des aspects ensemblistes ou éventuellement les axiomes de Kolmogorov (T1 t.a.r. 0.00641), ces derniers accompagnés en général d'un travail sur le registre symbolique (R3 t.a.r. 6.53e-007). Du groupe optimal d'individus typiques conformé par 45 exercices (un peu moins de la moitié du manuel) nous en transcrivons un (Exercice 65, Belin S, page 212) :

« Soit A, B et C trois événements d'un ensemble Ω sur lequel est définie une loi de probabilité P.

a. Calculer les probabilités P(A) et P(B) sachant que :

$P(A \cap \overline{B}) = 0{,}32, P(\overline{A} \cap B) = 0{,}20$ *et* $P(A \cup B) = 0{,}68$

b. Calculer P(C) sachant que $P(B \cup C) = 0{,}66$ *et* $P(B \cap C) = 0{,}10$

c. Calculer $P(A \cap B \cap C)$ *sachant que* $P(A \cap B) \cup (B \cap C) = 0{,}20$

d. Calculer $P(A \cup B \cup C)$ *sachant que* $P(A \cap C) = 0{,}14$ *»*

A la différence de l'exercice transcrit, ceux du groupe ont tendance à être courts, à poser moins de trois questions (L1 t.a.r 0.0232, Q1 t.a.r 0.0102) et à se placer dont les trois premiers quarts de ce manuel. Une grande vague de cette classe occupe le manuel de l'exercice 10 au 55, avec quelques autres dispersés au début et à la fin du chapitre.

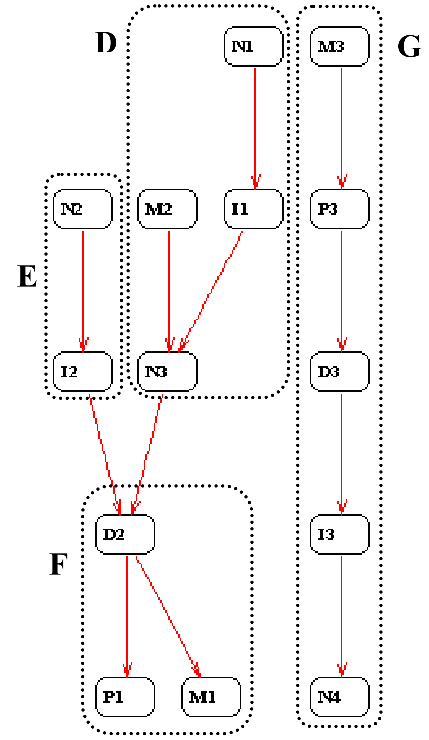
D'après cette première analyse de similarités, le manuel Belin de la filière S peut être caractérisé comme proposant trois types d'exercices :

- Un premier groupe fréquentiste. Faible en nombre, d'exercices plutôt longs où la notion fréquentiste de la probabilité est présente de manière explicite et abordée que par des simulations. Pour les exercices de cette classe, lorsqu'il y a un calcul de probabilité, il est courant que la fréquence d'apparition valide *a posteriori* ces calculs ; sinon et très rarement, ils se centrent sur la stabilisation de fréquence, estimant la probabilité à l'aide d'un échantillon.
- Un deuxième groupe bayésien. De la probabilité calculée, de présence plus prononcée, ils sont de nature bayésienne (épreuves génériques), mais l'interprétation n'est pas pertinente pour à la résolution du problème. Bien au contraire, ces exercices demandent de calculer explicitement des probabilités sans un retour sur la sémantique du calcul réalisé.
- Un troisième groupe d'exercices dépourvus de contexte. Nombreux, ils demandent d'opérer dans une abstraction en s'appuyant sur des relations logico mathématiques pour trouver leurs solutions (axiomes, combinatoire, etc.). Ce genre d'exercices ne fournit aucune trace permettant d'associer une interprétation au calcul de probabilité.

Ce premier ensemble de conclusions sera confirmé et complété par une étude du type implicatif.

Graphe implicatif

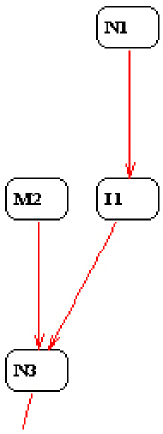
Graphe implicatif (seuil 0,85). Manuel Belin S



Graphique 19

Groupe D

Ce groupe implicatif concerne quatre des cinq variables du Groupe A de l'analyse des similarités que nous avons caractérisé comme fréquentiste (Tableau 22). D'une part ces variables confirment les conclusions tirées lors de l'analyse de similarités, et d'autre part elles enrichissent l'analyse par des relations implicatives relativement fortes.



Cod	Var	Mod	Card
M2	Manip	Simul	15
I1	Inter P	Oui	15
N3	Nat A	Amb	18
N1	Nat A	Série	4

Tableau 22

La règle N1 $\Rightarrow$ I1, qui d'ailleurs n'apparaît pas à des niveaux plus restrictifs de seuil, représente la relation conditionnelle *si la nature de l'objet est la série, alors il y a une interprétation de la probabilité*. Ceci nous indique que ce manuel traite une interprétation de la probabilité lorsque l'exercice est de nature fréquentiste. Ces exercices sont très peu nombreux dans l'ensemble du chapitre Probabilités du manuel Belin S (Card $_{N1}(I1) = 3/4$).

La règle I1 $\Rightarrow$ N3 présente déjà au seuil 0,99, pourrait s'exprimer sous la forme : *s'il y a une interprétation de la probabilité alors la nature de l'objet est ambiguë (*Card $_{I1}(N3) = 12/15$). La variable N3 nous indique de la concurrence de deux types de nature d'objets (générique et série). Cela nous indique du double rôle de la stabilisation de fréquences, d'une part pour les séries infinies (implication N1 $\Rightarrow$ I1) et d'autre part pour valider *a posteriori* le calcul d'une probabilité portée sur une épreuve générique par la stabilisation de fréquences (M2). Un exemple de cette utilisation a été donné dans l'exercice 78 (Groupe A). L'exercice suivant en est un autre (exercice 79, Belin S, page 214) :

> *On choisit au hasard un nombre entier compris entre 1 et 25 inclus et on détermine la somme de ses chiffres.*
> *1. a. Dresser la table de la loi de probabilité ainsi définie.*
> *b. Calculer son espérance et son écart type.*
> *2. b. Simuler 200 fois l'expérience aléatoire à l'aide d'une calculatrice.*
> *b. Dresser la table de fréquences à 0,01 des différentes sommes que l'on a obtenues.*
> *c. Calculer la moyenne et l'écart type de la série. Comparer les résultats à ceux de la question 1. b.*

L'exercice commence par préciser l'épreuve générique ; dans un deuxième temps il demande d'assigner les probabilités respéctives, continuant par un calcul de l'espérance et de l'écart type. Une fois ces calculs réalisés, il les reprend par une autre voie, celle d'une simulation ayant probablement la fonction de validation *a posteriori*.

Cette pratique, récurrente dans ce genre d'exercices, consiste à interroger sur une épreuve ponctuelle et à valider par l'ensemble. Dans aucun exercice nous n'avons trouvé une autre démarche, telle que par exemple interroger sur une épreuve ponctuelle et valider sur la même épreuve. Ceci demanderait d'une part la présence matérielle des objets, et d'autre part probablement de reconsidérer les critères d'argumentation et de preuve appris en classe de mathématiques. En effet, l'argumentation menant à probabiliser sur le résultat d'un prochain lancer d'une pièce de monnaie ne peut être invalidée par la non verification du résultat pronostiqué comme le plus probable.

La validation d'une probabilité du type épreuve générique se fait par un ensemble d'arguments considérés *a priori*, en état épistémique d'incertitude, et non pas en validant par le dévoilement de l'incertitude. Les exercices de ces manuels ne traitent pas ce genre de questions.

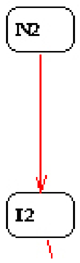
Dans ce sens, l'évocation du contexte contribue à contourner cette problématique. En mettant en état virtuel l'épreuve, on bloque tout intérêt possible de validation et de débats concernant la question. Dans ce sens, l'évocation des contextes n'a pas seulement une fonction de gestion du temps didactique, comme il semblait apparaître à première vue. Elle semble en avoir d'autres, comme celle d'encadrer les raisonnements, en favorisant les uns et bloquant les autres. Nous reviendrons sur ce sujet plus tard.

La deuxième implication (M2 $\Rightarrow$ N3), également présente au seuil de 0,99, exprime une autre quasi-implication intéressante : *si la manipulation est simulée alors l'objet est de nature ambiguë*. Les règles dans lesquelles interviennent la variable N3 (Nature ambiguë) deviennent des indices des difficultés à rester dans l'approche strictement fréquentiste. Les contextes glissent vers la notion bayésienne, généralement par l'intrication de signifiés (principe fréquentiste).

L'analyse implicative effectuée sur les exercices du manuel Belin S confirme le Groupe A fréquentiste trouvé par l'analyse des similarités. Néanmoins, et à cause de l'intrication de signifiés, ce groupe partage des règles avec le groupe bayésien E. Ces relations quasi-implicatives communes aux deux, nous les avons retenues dans le groupe F qui réunit tous les exercices fréquentistes et bayésiens. De cette manière, les classes D et E représentent les différences entre les exercices fréquentistes et bayésiens et la classe F retient leurs similitudes.

Groupe E

La règle N2 $\Rightarrow$ I2, déjà présente à un seuil de 0,99, symbolise la relation conditionnelle *si la nature de l'objet est du type générique, alors la probabilité calculée ne sera pas interprétée*. Cette quasi-implication concerne deux variables touchant un nombre important d'exercices (Tableau 23). En fait, en filtrant la base de données nous constatons qu'aucun exercice du type épreuve générique (strictement) n'a été interprété (Card($_{N2}(I2) = 36/36$), ce qui la transforme directement en une implication pour ce manuel.



Cod	Var	Mod	Card
N2	Nat A	Génér	36
I2	Inter P	Non	44

Tableau 23

Les exercices du type épreuve générique seront l'objet des calculs ou d'entraînement. Les calculs de probabilité ne deviendront pas des éléments guidant une prise de décision (I2$\Rightarrow$D2, Card$_{I2}(D2) = 42/44$) (Graphique 19). Dans ce sens, ni les épreuves génériques qui en général ne sont pas interprétées (N2$\Rightarrow$D2), ni les cas ambigus (N3), ne conduiront à des prises de décision.

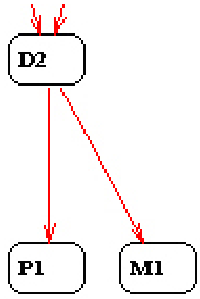
En fait, le lien I2$\Rightarrow$D2, présent au seuil 0,99 (s'il n'y a pas d'interprétation, alors il n'y a pas de décision), nous semble être un lien de nécessité. En effet, il serait impossible de fonder une décision où la probabilité soit concernée sans que celle-ci ne puisse assumer une interprétation. En d'autres termes, pour décider sur la base d'une probabilité il serait nécessaire qu'elle soit considérée dans sa dimension interprétative. Cette conclusion sera reprise dans nos expérimentations à partir d'un point de vue didactique, afin de créer les conditions pour l'émergence en classe des interprétations de la probabilité.

Notons en plus que tant une décision qu'une interprétation de la probabilité se développent sur le même type de registre, celui de la langue naturelle, orale ou écrite. Nous reviendrons plus tard sur ce sujet.

Le groupe bayésien E, comme tous d'ailleurs, nous a fourni quelques règles, d'une part nous renseignant sur les exercices, d'autre part nous outillant didactiquement pour nos expérimentations. Nous continuons par des règles partagées par les exercices tant bayésiens que fréquentistes.

Groupe F

Cette classe nous informe de caractéristiques communes aux exercices des deux approches. A la variable D2 (situations incertaines sans prise de décision) convergent les groupes D (fréquentiste) et E (bayésien) par les règles I2 ⇒ D2 et N3 ⇒ D2. Cette variable concerne donc toutes sortes de situations incertaines, nos données nous indiquent que pour ce manuel aucun exercice ne propose une prise de décision (Card D1=0). Plus de la moitié (60/104) décrivent des contextes incertains, dont une grande majorité (54/60) demande le calcul explicite de la probabilité (P1).



Cod	**Var**	**Mod**	**Card**
D2	Décis	Non	60
P1	Calc P	Expl	65
M1	Manip	Evoq	62

Tableau 24

La règle D2 ⇒ P1 présente déjà au seuil de 0,99, concerne des objets de toutes sortes de nature, soient ils la série, une épreuve générique ou les ambigus. Cette quasi-implication nous renseigne sur la probabilité dans sa dimension décisionnelle ; elle représente sémantiquement l'expression *s'il n'y a pas de décision à entreprendre, alors le calcul de la probabilité se fait explicitement*. Nous ne devons pas interpréter cette règle comme une relation de causalité dans laquelle la raison d'une demande explicite du calcul de la probabilité est due à la non considération de sa dimension décisionnelle. Ceci conduirait erronément à penser que le seul moyen de mobiliser un calcul de probabilité (autre que son explicitation) serait donc un contexte décisionnel. D'autres fonctions que la décisionnelle nous semblent envisageables, bien que nous nous en intéressons fortement. En tout cas, cette règle nous parle d'une possibilité absente pour la probabilité, celle de la considérer comme un outil décisionnel dans un contexte d'incertitude.

La règle D2 ⇒ M1 (s'il n'y a pas de décision à entreprendre, alors le contexte est du type évoqué) ne nous semble pas en soi d'intérêt sémantique, sinon par le fait que les deux variables semblent une conséquence du fort intérêt pour les aspects calculatoires au détriment des interprétatifs.

Cette classe commune aux deux approches interprétatives nous indique finalement qu'en amont de toute sorte d'interprétations possible, la probabilité dans le manuel Belin S ne parvient pas à devenir un élément rationnel de prise de décision.

Groupe G

Le dernier groupe du graphe implicatif du manuel Belin S révèle un enchaînement de règles, M3 $\Rightarrow$ P3 $\Rightarrow$ D3 $\Rightarrow$ I3 $\Rightarrow$ N4. Cet ensemble de quasi-implications ne constitue pas un apprentissage pertinent, son enchaînement n'est qu'une conséquence logique de leur sémantique. Ces variables caractérisent les exercices sans contexte (C5, variable secondaire). Si un exercice manque de contexte, il ne convient pas de s'interroger sur une éventuelle manipulation (M3), ni de se poser la question sur une plausible décision à prendre (D3), ni non plus, finalement, de s'intéresser à une interprétation de la probabilité (N4).

Cod	Var	Mod	Card
M3	Manip	Non C	27
P3	Calc P	Non C	34
D3	Décis	Non C	44
I3	Inter P	Non C	45
N4	Nat A	Non C	46

Tableau 25

Ce groupe reprend en règles implicatives les variables ressemblées par leurs similarités dans le groupe C. Dans les deux groupes ces variables restent isolées des autres, cette distanciation du reste des exercices nous indique d'un groupe solidement défini dont sa principal caractéristique concerne la focalisation sur des aspects algébriques ou numériques, sans aucune prise en compte d'un contexte incertain.

De l'étude implicative du manuel Belin S nous retenons des conclusions dans deux directions, d'une part l'existence de contextes tant bayésiens que fréquentistes, et d'autre part la tendance de ce manuel à traiter une dualité qui ne parvient pas à être contournée. Voici les trois profils d'exercices repérés par les outils du logiciel CHIC :

- Un groupe frequentiste : Groupe A (analyse de similarités) ←→ Groupe D (graphe implicatif)
- Un groupe bayésien : Groupe B (analyse de similarités) ←→ Groupe E (graphe implicatif)
- Un groupe d'exercices sans contexte : Groupe C (analyse de similarités) ←→ Groupe G (graphe implicatif)

- Des caractéristiques communes aux exercices fréquentistes et bayésiens mises en évidence par la confluence des groupes D et E vers le groupe F.

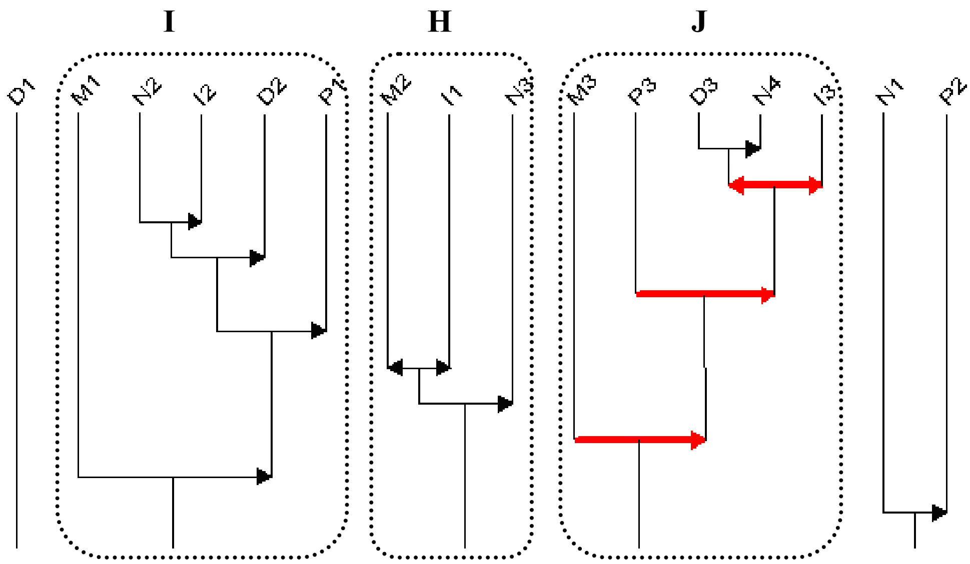
Pour ce qui concerne les tendances, elles nous renseignent sur les moyens utilisés par ce manuel pour traiter la dualité de la probabilité. Par exemple, quelques quasi-implications nous apprennent que ce manuel résout le conflit en contournant le traitement de la notion bayésienne mais non ses situations, laissant entièrement à la charge de l'élève la construction de cette notion :

- La règle (déjà présente à un seuil de 0,99 (Groupe E, N2 $\Rightarrow$ I2)) : *si la nature de l'objet est du type générique (bayésienne), alors la probabilité calculée ne sera pas interprétée.*
- L'implication du Groupe D (N1 $\Rightarrow$ I1) : *si la nature de l'objet est la série alors il y a une interprétation de la probabilité.*

Nous continuons notre analyse par l'arbre cohésitif, à la recherche cette fois-ci de possibles méta règles.

Arbre cohésitif

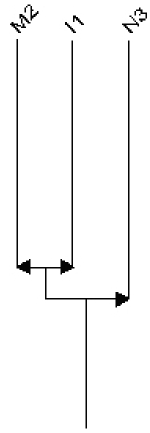
Arbre cohesitif. Manuel Belin S



Graphique 21

Groupe H

Toutes les variables de ce groupe sont présentes dans les groupes fréquentistes de l'arbre de similarités et du graphe implicatif (Groupe D).



Cod	**Var**	**Mod**	**Card**
M2	Manip	Simul	15
I1	Inter P	Oui	15
N3	Nat A	Amb	18

Tableau 26

Ce groupe partage avec le Groupe A presque tous les individus typiques (11/12) et toutes les variables secondaires typiques. Pour le Groupe H, il s'agit d'exercices longs (L2 t.a.r. 0,11) posant plus de quatre questions (Q3 t.a.r. 0,00681) avec une forte présence du tableur (T3 t.a.r.0,00804), mettant en place de situations de cas élémentaires équiprobables (H1 t.a.r. 0,0657).

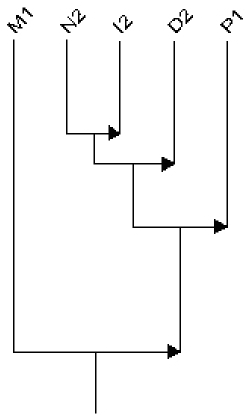
Ce groupe confirme donc l'existence d'une classe constituée par des exercices de type fréquentiste. Ainsi les Groupe H et A seraient des vues complémentaires d'un même ensemble (fréquentistes). Au profil déjà décrit par le Groupe A, nous joignons les méta règles du graphe cohésitif du Groupe H :

M2 $\Longleftrightarrow$ I1
(M2 $\Longleftrightarrow$ I1) $\Rightarrow$ N3

La double implication (M2 $\Longleftrightarrow$ I1) ($\text{Card}_{M2}(I1) = 12/15 = \text{Card}_{I1}(M2) = 12/15$) met en évidence le choix du manuel Belin S de s'assurer de ne pas proposer d'exercices concernant une interprétation par la seule voie d'une simulation. La pseudo-réalité de la simulation cherchant toujours à traiter la stabilisation de fréquences et sa valeur de convergence. Néanmoins, rappelons que le manuel Belin S fait intervenir cette notion dans ses deux faces, en tant que notion fréquentiste proprement dite et en tant que raison à croire pour évaluer une épreuve générique. La variable N3 (nature ambiguë) est un indicateur de cette utilisation de la notion fréquentiste.

Groupe I

Pour ce groupe, non seulement les variables concernées sont les mêmes que celles du groupe B de l'arbre de similarités (bayésien), mais les groupes optimaux d'individus typiques sont identiques, et la hiérarchisation cohésitive reproduit dans l'ordre le commencement de l'arbre de similarités ((N2⇒I2) ⇒ D2 ; (N2, I2), D2).



Cod	Var	Mod	Card
I2	Inter P	Non	44
N2	Nat A	Génér	36
D2	Décis	Non	60
M1	Manip	Evoq	62
P1	Calc P	Expl	65

Tableau 27

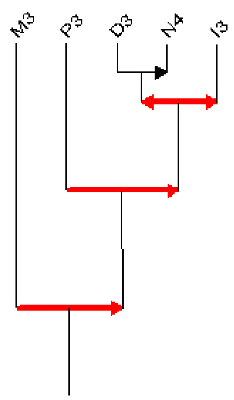
Pour la règle N2⇒I2 nous avons un nombre non négligeable d'exercices (Card $_{N2}(I2) = 36/36$), tous de nature épreuve générique pour lesquels la probabilité n'est pas interprétée. Pour ces exercices, les calculs sollicités dans l'énoncé ne deviennent pas des outils décisionnels ((N2⇒I2) ⇒ D2). Cette métarègle nous informe qu'aucun exercice ne demand de faire ce qu'à nos yeux il semble impossible d'effectuer: solliciter une décision sans que l'interprétation de la probabilité ne soit pas considérée. En effet, la règle I2⇒D2 nous semble exprimer une condition de nécessité : pour entamer un argument décisionnel, il faut que la probabilité concernée s'habille d'un de ses signifiés ; sans cette interprétation, une décision la concernant devient incomplète et inconsistante.

Ce groupe bayésien met en évidence, par P1, une autre caractéristique qui nous renseigne sur l'intention de ce manuel de se focaliser sur des aspects calculatoires de la probabilité, en particulier par les registres algébrique et numérique (((N2⇒I2) ⇒ D2) ⇒ P1).

Groupe J

A l'intérieur de ce groupe nous trouvons les exercices les plus courts en nombre de lignes. Ils se caractérisent par le fait d'utiliser des propriétés strictement mathématiques de la probabilité. Ces exercices sembleraient faire partie d'une pratique assez répandue selon laquelle un ensemble d'exercices est consacré à des aspects exclusivement mathématiques en économisant la description d'un contexte incertain ; puis, une fois les aspects mathématiques relativement assurés, un autre ensemble d'exercices est proposé pour traiter des aspects sémantiques ou d'application de la probabilité.

Un indice de cette progression peut être trouvé dans les exercices du groupe optimal de cette classe, ils ont tendance à être placés au début du chapitre. D'ailleurs, si cette organisation séquentielle se reproduit en classe, et vu les contraintes de temps, l'enseignement de relations logicomathématiques serait traité en priorité et puis, dans un deuxième temps, les aspects plus liés à la sémantique de la probabilité.



Cod	Var	Mod	Card
M3	Manip	Non C	27
P3	Calc P	Non C	34
D3	Décis	Non C	44
I3	Inter P	Non C	45
N4	Nat A	Non C	46

Tableau 28

Cette séquence de séparation entre les aspects mathématiques d'une part et les sémantiques d'autre part, pourrait être un autre facteur qui contribue à reléguer le traitement des interprétations de la probabilité. Le pilotage du temps consacré à un thème s'accentue à la fin du chapitre, et c'est là où se placeraient les exercices où les interprétations ont plus de place, les coupures de dernière minute se répercuteraient sur ces exercices plutôt que sur les techniques, diminuant encore plus leurs chances d'être traités en classe.

L'analyse du manuel Belin de la filière scientifique nous a permis de dessiner trois grands groupes d'exercices pour le chapitre Probabilité. Nous résumons chacun d'eux :

Groupe fréquentiste

Ce groupe est présent dans les trois analyses effectuées : Groupe A (Analyse de similarités), Groupes D–F (Analyse implicative) et Groupe H (Analyse cohésitive).

En termes de pourcentage, ces exercices ne dépassent pas 12 % des 104 du chapitre Probabilité. Ce sont en général les plus longs et ceux qui posent les plus de questions ; cela semblerait être la conséquence de deux facteurs. Le premier concerne la construction d'un contexte d'attente à long terme. En effet, l'explicitation de la reproduction de l'expérimentation sous les mêmes conditions est une tâche qui demande une attention considérable. Le deuxième concerne le choix du manuel de substituer aux expérimentations réelles des situations simulées. Le travail de spécification de l'environnement informatique alourdit la taille des exercices.

Les tirages simulés sont sollicités afin de s'effectuer sur un ordinateur ou sur une calculatrice. Les modèles sousjacents à ces données sont toujours ceux d'événements élémentaires équiprobables. Cela est une autre caractéristique de cette classe fréquentiste, on ne simule que des situations dont la stabilisation de fréquences peut être déduit *a priori* par l'admission d'hypothèses d'équipossibilité. C'est ainsi qu'on ne trouve pas de contextes où la probabilité puisse s'estimer que par la convergence de la proportion d'apparition de l'événement.

La stabilisation des fréquences dans des contextes de cas élémentaires équiprobables est utilisée pour boucler les calculs de la probabilité *a priori*. Elle remplit la fonction de validation par l'estimation *a posteriori*. Dans ce sens, la logique déterministe semble s'installer dans le chapitre Probabilités, mêmes si les situations sont incertaines. Les probabilités sur une épreuve générique évaluées par le principe fréquentiste sont traitées comme étant vraies (ou fausses) et plausibles d'être confirmées (ou réfutées). En fait, l'aspect indéterministe de la situation n'est considéré que comme une condition contextuelle au concept de probabilité. Une fois calculées les probabilités, le contexte indéterministe n'intervient plus, l'élève n'est pas invité à utiliser ses productions sur la situation incertaine, en particulier, aucune prise de décision n'est demandée.

Quelques exercices de ce groupe associent d'une manière ambiguë le calcul de la probabilité à la stabilisation de fréquences, rares sont ceux qui sont nettement de nature fréquentiste, un nombre important fait intervenir les objets de deux natures différentes, le calcul est sollicité pour une épreuve générique puis la situation pivote sur la série pour estimer la stabilisation de fréquences, laquelle revient sur l'épreuve générique pour valider le calcul.

Groupe bayésien

Cette classe se manifeste elle aussi avec les trois outils statistiques utilisés : Groupe B (Analyse de similarités), Groupe E-F (Analyse implicative) et Groupe I (Analyse cohésitive).

Ce groupe est sensiblement plus nombreux que le fréquentiste. Les exercices sont plutôt longs mais toujours plus courts que ces derniers. Ils posent un peu moins de questions, entre 3 et 4 en général. L'évaluation des épreuves génériques caractérisant cette classe se fait sur la base de cas élémentaires équiprobables. Ces exercices définissent les objets de cette nature d'une manière tacite. En fait, l'intérêt étant de préciser les cas favorables et les possibles et non pas l'objet en soi. Cet objet bayésien est bien entendu non intentionné, de toutes manières, sa nature sera insignifiante aux fins de ces exercices qui ne demandent que l'évaluation de la probabilité.

Ces exercices ne reviennent pas *a posteriori* sur une valeur de probabilité calculée *a priori*. En réalité, pour le paradigme bayésien, une telle démarche manque de sens. Par exemple, (Exercice 56, Belin S, page 211, transcrit en Groupe B) : *« On a tiré une carte dont on ne connaît pas la couleur ; on demande la probabilité que la carte tirée soit rouge »*, en se basant sur le principe d'équipossibilité on l'évalue numériquement à 1/8. Dans ce cas, il n'y aurait pas de sens de s'interroger sur une possible correction de cette évaluation ni qu'on s'est trompé non plus, si on constate *a posteriori* que la carte était effectivement rouge, tout simplement parce que la probabilité portée sur cette épreuve n'est qu'une mesure que l'on estime *a priori*. Une bonne partie des exercices de ce manuel concerne des épreuves génériques. Néanmoins la question de critères autres que celui d'équiprobabilité est contournée par un dispositif consistant à, premièrement, ne pas formuler des questions concernant sont interprétation, deuxièmement, toujours évoquer l'épreuve et dernièrement, ne traiter que celles qui sont susceptibles de reproduction. L'articulation de ces trois conditions laisse toujours, le cas échéant, la possibilité d'une reformulation en termes fréquentistes. En effet, ne pas s'interroger sur le signifié de l'objet probabilité est un élément important, le manque constant d'intérêt pour la sémantique des calculs effectués pourrait s'installer dans le contrat didactique de la classe en bloquant tout éventuel intérêt pour l'interprétation. L'absence de l'objet (évocation) permet, si les circonstances le requièrent, une rapide reconversion de la situation vers un contexte fréquentiste. Enfin, la plausibilité de reproduction facilite que cette reconstruction soit de type fréquentiste.

Groupe sans contexte

Ce groupe s'est aussi manifesté ors des trois analyses effectués avec le logiciel CHIC : Groupe C (Analyse de similarités), Groupe G (Analyse implicative et Groupe J (Analyse cohésitive).

Les exercices decontextualisés échappent à toute tentative d'interprétation de la probabilité. Quelques-uns n'arrivent pas à traiter une probabilité restant sur des concepts plus élémentaires, mais une grande majorité est consacrée à des tâches calculatoires dont les transformations plausibles s'effectuent par propriétés ensemblistes ou axiomes de Kolmogorov.

Ces exercices ont une présence importante dans ce manuel de la filière Scientifique, ils ont tendance à être courts, posant peu de questions. Ils se focalisent sur des notions plutôt ciblées, de caractère bien évidemment abstrait.

Ce manuel dédie un espace important à ce genre d'exercices, ils occupent à peu près la moitié du manuel et sont placés plutôt au début de chapitre. En fait, si nous réunissons tous les exercices pour lesquels aucune interprétation n'est concernée (groupe bayésien et groupe sans contexte) nous observons une forte tendance à se concentrer sur des aspects calculatoires (89/104 exercices, 85 %). De leur part, les exercices fréquentistes, même longs et placés plutôt à la fin du chapitre, sont les seuls à donner une place à l'utilisation des TICE.

Bref, les interprétations de la probabilité sont traitées dans ce manuel différemment, si une interprétation est explicitée, elle sera fréquentiste, et cela,soit comme notion centrale lorsque l'objet est par nature la série, soit comme notion secondaire (principe fréquentiste). Dans ce dernier cas, son vrai rôle n'est pas dévoilé. Pour les épreuves génériques qui évaluent la probabilité par le principe d'équipossibilité, elles restent sans interprétation, l'attention se focalise sur des tâches calculatoires. Nous analyserons un autre manuel de la filière Scientifique, celui de la maison Nathan, édité aussi l'année 2001.

2.6 Manuel Nathan S

Ce manuel édité pour la maison Nathan distribue les 73 exercices du chapitre Probabilité en quatre sections : *Maîtriser le cours*, *Pour apprendre à chercher*, *Pour progresser* et *Problèmes de synthèse*. Le tableau ci-contre montre leur distribution par modalité.

Distribution du nombre d'exercices du manuel Nathan S par modalité

Lignes		Questions			Représentations				Tâches				Contexte				
L1	L2	Q1	Q2	Q3	R1	R2	R3	R4	T1	T2	T3	T4	C1	C2	C3	C4	C5
53	20	32	20	21	7	4	15	47	54	17	0	2	48	12	4	4	5

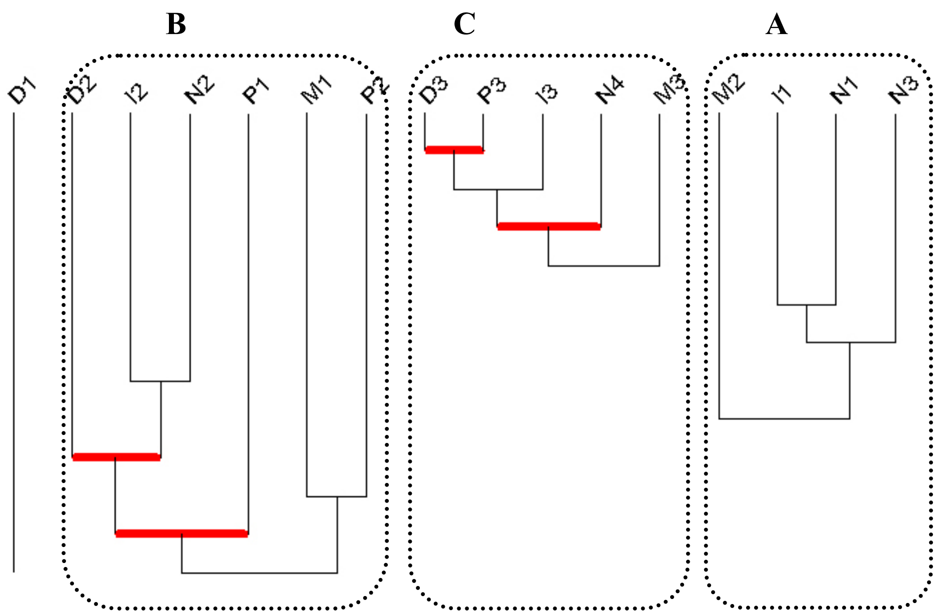
Décision			Hypothèses				Manipulation			Calcul P			Interprétat P			Nature A			
D1	D2	D3	H1	H2	H3	H4	M 1	M 2	M 3	P1	P2	P3	I1	I2	I3	N1	N2	N3	N4
0	62	11	68	4	0	1	66	2	5	64	2	7	6	56	11	1	39	22	11

Pour notre analyse du manuel Nathan S ainsi que pour les deux autres de la filière ES nous allons utiliser les même désignations (A, B, C, …) que celles des groupes des analyses du manuel Belin S, et cela si les profils se correspondent, cas contraire nous nous servirons d'autres lettres de l'abécédaire.

Arbre de similarités

L'arbre de similarités présente trois grands groupes assez proches à ceux du manuel Belin S. Une des principales différences consiste en le déplacement de la variable P2 (Calcule de P implicite) de la classe fréquentiste de Belin S vers la classe bayésienne de Nathan S. Nous analyserons en détail ces trois regroupements.

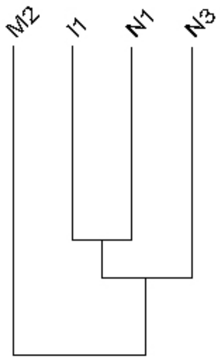
Arbre de similarités. Manuel Nathan S



Graphique 23

Groupe A

Cette première classe est faible en nombre, I1 joint N1 par très peu d'exercices (Card($I1 \cap N1) = 1$). A cet ensemble se rend N3 formant un groupe caractérisé par des problèmes dont l'interprétation intervient dans l'énoncé et les objets sont de nature fréquentiste ou ambiguë avec quelques rares simulations. L'indice de similarité de la sous-classe ((I1, N1) N3) reste assez élevé, même celui de la classe générale après l'arrivée de M2 (0.942196).



Cod	Var	Mod	Card
M2	Manip	Simul	2
I1	Inter P	Oui	6
N3	Nat A	Amb	22
N1	Nat A	Série	1

Tableau 29

D'ailleurs, la sous-classe ((I1,N1) N3) possède le même groupe optimal d'individus typiques que la classe générale (((I1,N1) N3) M2). Ce groupe optimal est constitué en fait par un seul exercice, nous le transcrivons (Exercice 29, Nathan S, page 228) :

« Le jeux de « passe-dix » consiste à jeter trois dés, on gagne si la somme des points obtenus dépasse 10. Le chevalier de Méré constatait qu'en pratique on

gagnait plus souvent avec 11 qu'avec 12. Cela contredisait son raisonnement, que voici :

« Il y a six possibilités de marquer 11 points :{4,4,3}, {5,3,3}, {5,4,2}, {5,5,1}, {6,3,2}, {6,4,1} et six possibilités de marque 12 points : {4,4,4}, {5,4,3}, {5,5,2}, {6,3,3}, {6,4,2}, {6,5,1}. Donc la probabilité de marquer 11 est égale à celle de marquer 12. »

L'erreur du chevalier est de s'en tenir aux issues observables et de les croire équiprobables. Or si l'on veut des issues équiprobables il faut distinguer les dés.

1. *Vérifiez qu'en distinguant les dés, il y a, par exemple, six façons d'obtenir 5, 4 et 3.*
2. *Vérifiez qu'il y a vingt-cinq façons d'obtenir 12 et vingt-sept d'obtenir 11.*
3. *Prouvez que la probabilité d'obtenir 12 est 25/216 et celle d'obtenir 11 est 27/2176 ».*

Pour cet exercice, la nature de l'objet à probabiliser est la série (N1), la demande de son calcul est explicite (P1), et ce dernier est associé à une interprétation (I1). Bien qu'il constitue tout seul la classe optimale, il ne l'est pas en la contribution au groupe. Pour la première sous-classe, le groupe optimal de la contribution des individus est conformé par 6 exercices, pour la classe générale, deux autres se joignent. C'est par ces exercices qui la modalité N3 apparaît (nature ambiguë) dans l'arbre de similarités et non pas par l'exercice optimal.

Le manuel Nathan S, comme l'a fait Belin S, demande lui aussi de calculer la probabilité sur des objets de nature ambiguë (N3, Nathan S 22/73 exercices ; Belin 18/104). Cela signifie que, soit il y a une imprécision sur la nature de l'objet, soit il y a présence d'objets de nature différentes pour un même calcul de probabilité. Pour le premier cas, le plus courant dans ce manuel, il s'agit en général d'une imprécision par une énonciation formelle sans traces de sa sémantique (« déterminer la loi de probabilité pour l'événement A…l'événement A : tirer un quatre »). Pour le deuxième cas, et comme nous l'avons précisé pour le manuel Belin S, la présence d'objets de nature différente se produit lorsque l'exercice demande de calculer la probabilité sur une épreuve générique par le principe d'équipossibilité et puis de la confronter à la fréquence d'apparition de l'événement (principe fréquentiste).

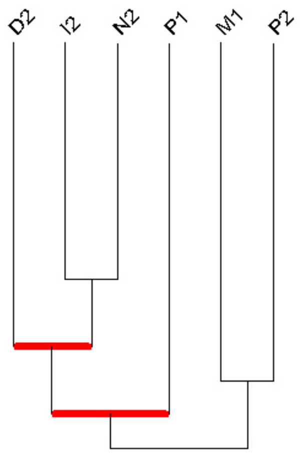
Finalement, cette classe fréquentiste se caractérise par des exercices longs posant plus de 4 questions, mais moins nettement que pour l'autre manuel de la même filière (L2 t.a.r 0,241 et Q3 t.a.r. 0,251).

Nous trouvons donc pour ce manuel, une classe fréquentiste de caractéristiques similaires à celle de l'autre manuel de la même filière, cette ressemblance concerne tant les variables principales que les secondaires.

Groupe B

La classe B, bayésienne, un peu moins homogène que son homonyme du manuel Belin S commence le rassemblement par les mêmes variables ((I2, N2) D2), dont la principale différence se manifeste en la présence de la variable P2 qui s'est déplacée de la classe fréquentiste chez Belin S vers la bayésienne chez Nathan S.

Ce groupe B dont la probabilité calculée explicitement (P1) n'est pas l'objet d'une interprétation (I2) ni d'une prise de décision (D2), se caractérise en plus par des contextes de type épreuve générique (N2). Le cardinal des variables concernées (Tableau 30) indique un nombre important d'exercices (Card($I2 \cap N2 \cap D2$)= 39, Card($I2 \cap N2 \cap D2 \cap P1$)=38). En fait, plus de la moitié sont des épreuves génériques pour lesquelles aucune interprétation ne sera requise dans l'énoncé.



Cod	Var	Mod	Card
D2	Décis	Non	62
I2	Inter P	Non	56
N2	Nat A	Génér	39
P1	Calc P	Expl	64
M1	Manip	Evoqu	66
P2	Calc P	Implic	2

Tableau 30

Un exercice optimal typique de cette classe illustre quelques-unes de ces variables (Exercice 60, Nathan S, page 232) :

« Un élève répond au hasard à un Q.C.M. qui comporte cinq questions. A chaque question, deux réponses sont proposées, dont une seule est correcte. Le correcteur attribue quatre points pour une réponse exacte, mais enlève deux points pour une réponse fausse. Lors de la note finale, il porte 0 pour

tout total négatif ou nul. Ainsi la note finale N est une variable aléatoire définie sur l'ensemble de toutes les réponses équiprobables au Q.C.M. Quelle note moyenne peut espérer cet élève ? »

L'objet à probabiliser est ici une épreuve générique (N2), le supposé élève s'intéresserait à son examen, et si à sa place on fait recours à ce qu'il arrive fréquemment ce n'est que pour l'utiliser comme un argument, une raison pour évaluer numériquement les chances de réussir l'examen. Un professeur pourrait prendre une approche fréquentiste lorsque il considère la totalité d'examens, mais pas l'élève qui souhaite réussir le sien.

Néanmoins, le principe fréquentiste ne semble pas être le recherché pour le manuel. L'indice d'un autre est glissé dans l'énoncé (« équiprobables »). En fin, les deux principes sont toujours plausibles de devenir l'argument pour évaluer numériquement la probabilité, vu que tous les exercices se basent sur le principe d'équipossibilité.

Cet exercice présente deux singularités, premièrement le calcul de la probabilité se demande indirectement (Card (P2) = 2), deuxièmement cette demande ne se fait pas par l'objet mathématique (espérance mathématique) comme d'habitude, mais plutôt par sa sémantique, bien qu'un indice est glissé par le mot « *espérer* » et qu'une partie du travail de modélisation est déjà réalisé (« *Ainsi la note finale N est une variable aléatoire définie sur l'ensemble de toutes les réponses équiprobables au Q.C.M »).* Cet exercice est un des rares exemples où même si timidement un énoncé parcourt le sens *signifié* → *objet mathématique*. En général, si ces deux éléments sont présents, ils sont reliés dans la direction opposée *(objet mathématique* → *signifié).* D'abord on demande de calculer la probabilité et puis, très rarement (Card (I2) = 6) ce calcul est soumis à une interprétation.

D'ailleurs, cette classe bayésienne possède un groupe optimal d'éléments typiques beaucoup plus nombreux que la fréquentiste (Groupe B : 39 exercices ; Groupe A : 1 exercice). Les variables secondaires typiques nous renseignent sur des exercices posant assez de questions (Q2 et Q3 t.a.r. 0,102 et 0,159 respectivement) et de situations qui se déroulent en contextes de la vie quotidienne ou concernant le monde du travail (C2, t.a.r 0,0062) ce qui rend le manque de décision (D2) plus regrettable vue leur transcendance par rapport aux exercices ludiques.

Pour l'essentiel, les mêmes caractéristiques se reproduisent dans cette classe bayésienne dans les deux manuels de la filière scientifique.

Groupe C

Cette classe chez Nathan S, est aussi très proche de son homonyme du manuel Belin S, elle représente les exercices dont les contextes sont absents ou les questions ne concernent pas la probabilité. Ce manuel s'engage moins sur cette voie que Belin S, les cardinaux de variables concernés pour Nathan S varient entre 5 et 11 sur un total de 73 exercices, tandis que pour Belin S entre 27 et 45 sur un total de 104 exercices.

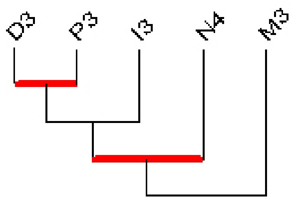


Tableau 31

Cod	Var	Mod	Card
D3	Décis	Non C	11
P3	Calc P	Non C	7
I3	Inter P	Non C	11
N4	Nat A	Non C	11
M3	Manip	Non C	5

Nous transcrivons un exercice de son groupe optimal d'individus typiques (cardinal 7) (Exercice 28, Nathan S page 228) :

« On lance deux dés équilibrés dont les faces sont numérotées de 1 à 6. Dans chacun des cas suivants indiquez su les événements A et B sont incompatibles.

1. A : « Les deux chiffres sont impairs », B : « la somme des chiffres est impaire »

2. A : « les deux chiffres ont même partie », B : « la somme des deux chiffres est un multiple de 3 ». »

Dans cet exercice ce n'est pas le manque de contexte qui bloque les caractéristiques qui nous intéressent d'analyser, mais plutôt l'absence de référence à la notion de la probabilité dans les questions de son énoncé.

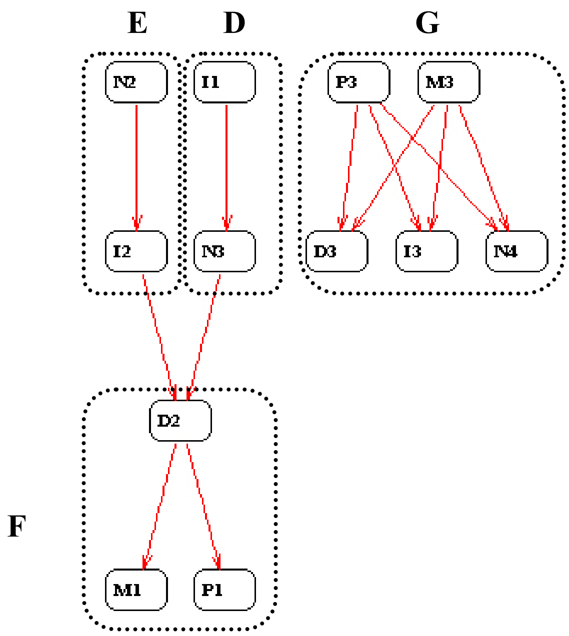
En corrélation avec le cardinal de variables concernées, le groupe optimal d'individus typiques est assez réduit (7 exercices). Placés tout au début du chapitre ou plus rarement au milieu, ces exercices semblent faire part des bases de ces groupes d'exercices hiérarchisés aux quels nous avons fait référence plus haut.

Les exercices de ce groupe, dédiés à des notions plutôt ciblées, se caractérisent par sa taille réduite (L1 t.a.r. 0,243), par le peu de questions posées (Q1 t.a.r. 0,188) et par des contextes ayant tendance à être ludiques (C1 t.a.r. 0,311).

L'analyse de similarités appliquées sur ces deux manuels nous renseigne donc sur des tendances proches, en amont de différences telles que l'utilisation de simulations pour le manuel Belin S et la moindre proportion d'exercices sans contexte que propose Nathan S. L'analyse implicative aidera à mieux cerner la description de ces trois groupes.

Graphe implicatif

Graphe implicatif (seuil 0,85) Manuel Nathan S

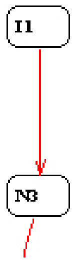


Graphique 25

Groupe D

Pour un seuil de 0,85, ce groupe retient une seule règle (I1⇒N3) exclusive à la notion fréquentiste, à la différence du manuel Belin S qui en affichait trois (N1⇒I1, I1⇒N3 et M2⇒N3). Les quasi-implications du Groupe F (D2⇒M1, D2⇒P1), sont partagées par les groupes tant fréquentiste (D) que bayésien (E), devenant ainsi des caractéristique communes à toute sorte d'exercices contextualisés dont le calcul de la probabilité a lieu (caractéristique qui par ailleurs a déjà été repérée chez le manuel Belin S).

Il faut descendre au niveau 0,60 du seuil pour que la variable N1 (Nature de l'objet : Série) fasse son apparition dans ce groupe avec l'implication (N1⇒II). Cette absence semble se devoir d'une part au faible nombre d'exercices proposés par Nathan S par rapport à Belin S (73 et 104 exercices respectivement), et d'autre part au très faible nombre d'exercices du manuel Nathan S dont la nature des objets est la série (Card(N1)=1).



Cod	Var	Mod	Card
I1	Inter P	Oui	6
N3	Nat A	Amb	22

Tableau 32

Toutefois, ce groupe répond aux caractéristiques fréquentistes, principalement par la présence de la variable N3 qui rend compte ici d'exercices dont le calcul de probabilités s'effectue sur une épreuve générique par le principe d'équipossibilité et puis confirmé par la stabilisation de fréquences. C'est ainsi donc que l'interprétation fréquentiste participe en tant que raison fondant l'évaluation numérique d'une probabilité portée sur un événement singulier mais plausible de reproduction.

La règle I1⇒N3 (qui apparaît pour la première fois à un seuil de 0,92) indique que le choix de Belin S n'est pas isolé. Les manuels de la filière scientifique font jouer le principe fréquentiste pour l'évaluation d'épreuves génériques. Malgré cet usage répandu, ce manuel ni l'autre ne dévoilent les vrais rôles de chacune des interprétations.

Ce groupe, duquel nous ne retenons que cette première quasi-implication, sera complété par celles du groupe F qui réunit à mode d'intersection, des règles communes aux groupes bayésien et fréquentiste.

Groupe E

Les cardinaux des variables concernées pour la seule règle (N2⇒I2) de ce groupe sont moins distanciés que ceux de l'implication du groupe D, et en même temps plus importants en nombre. En fait cette règle est déjà présente à un seuil de 0, 99 ; à ce seuil, elle n'est accompagnée que par quelques unes du groupe très serré d'exercices sans contexte et par l'implication I2⇒D2 qui ne fait que mettre en évidence le rapport de nécessité de l'interprétation pour une prise de décision.

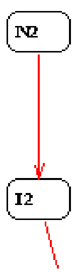


Tableau 33

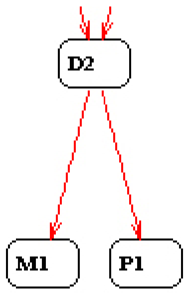
Cod	Var	Mod	Card
N2	Nat A	Génér	39
I2	Inter P	Non	56

La règle N2⇒I2, nous renseigne sur la place cachée de la notion bayésienne, en particulier sur le traitement donné aux épreuves génériques. Celles-ci sont largement présentes dans ce manuel, (Card (N2)=39), sur cette notion reposent plus de la moitié des exercices de ce manuel. Néanmoins, l'interprétation sous-jacente au calcul de la probabilité ne est relevante pour ces exercices (I2)

C'est ainsi que les exercices de nature bayésienne qui mobilisent le principe d'équipossibilité ne s'intéressent pas à l'interprétation de la probabilité. Les seuls à le faire s'assurent d'introduire une simulation pour que la notion fréquentiste soit la notion retenue (I1⇒N3). Finalement, très rares sont les exercices de nature fréquentiste stricto sensu (Card(N1)=1). Cela semble une conséquence du choix de ce manuel de ne pas proposer d'expérimentations réelles ni non plus de les remplacer par des simulations.

Groupe F

Ce groupe résume des caractéristiques communes aux deux précédents, vers D2 converge I2 et N3, la première des règles (*s'il n'y a pas d'interprétation alors aucune décision est entreprise avec le calcul de la probabilité réalisé)* informe d'une tendance dans ce manuel de ne pas s'engager dans la problématique de l'utilité de la probabilité en prises des décisions et concerne indirectement N3 par la double présence des objets de nature générique et fréquentiste.



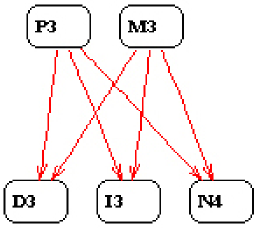
Cod	**Var**	**Mod**	**Card**
D2	**Décis**	**Non**	**62**
P1	**Calc P**	**Expl**	**64**
M1	**Manip**	**Evoq**	**66**

Tableau 34

D'ailleurs, et comme a été repéré chez Belin S, le manuel Nathan S s'occupe aussi de la probabilité à partir d'une approche plutôt logico déductive, les énoncés proposent de « prouver », « vérifier », etc. (exercice 29, Nathan S, page 228, transcrit en Groupe A, fréquentiste). Les propositions ainsi analysées ne dépassent pas le paradigme dichotomique du déterminisme. Tout se passe comme si les valeurs de probabilités ne seraient que correctes ou incorrectes, comme s'il serait toujours possible de déterminer si la valeur logique de la réponse est soit vraie soit fausse.

Groupe G

Ce groupe, identique au groupe G du manuel Belin S, reprend les variables de l'arbre de similarités réunies dans le groupe C. Dans cette classe se trouvent les exercices qui se caractérisent soit par un manque de contexte soit par un non-traitement du concept de probabilité, ce qui empêche toute analyse interprétative de la probabilité.



	Cod	Var	Mod	Card
	M3	Manip	Non C	5
	P3	Calc P	Non C	7
	D3	Décis	Non C	11
	I3	Inter P	Non C	11
Tableau 35	N4	Nat A	Non C	11

Rappelons que la qualité du rapprochement entre ces variables est en partie due au fait que quelques modalités sont une conséquence logique d'autres, ainsi, si un énoncé ne propose pas un contexte (C5), indéfectiblement il n'y aura pas de manipulation (M3) ni non plus d'interprétation possible (I3).

De toute manière, le Tableau 35 semble indiquer un pari moins important de la part du manuel Nathan S sur ce genre d'exercices, au moins par rapport à Belin S, même en termes de pourcentage, ces exercices sont significativement moins fréquents chez le premier.

L'analyse implicative sur le manuel Nathan S confirme nos hypothèses en deux directions, premièrement elle montre une approche dont la dualité d'interprétation n'est pas contournée par les situations proposées mais plutôt par les questions posées. Ces dernières se focalisent sur des aspects calculatoires lorsque l'objet est bayésien et sur les sémantiques lorsque la fréquence apparaît dans le contexte. Deuxièmement, cette voie semble se généraliser, le manuel Nathan S du même que Belin S, tous les deux résolvent le conflit par un blocage de l'interprétation de situations bayésiennes en le remplasant par des tâches calculatoires.

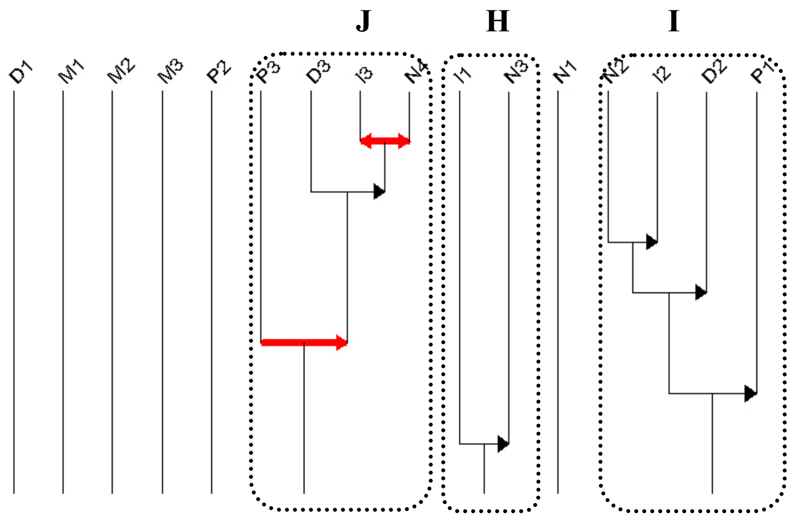
Nous continuerons par l'arbre cohésitif qui nous permettra de trouver des métarègles dans l'ensemble des exercices du manuel Nathan S.

Arbre cohésitif

Trois grands groupes se dessinent dans l'arbre cohésitif laissant isolées un nombre important de variables, les plus significatives sont :

- M1 et M2, manipulations évoquées et simulées respectivement. L'isolement de la première semblerait s'explique par son omniprésence, ce qui anéantirait l'étonnement de la trouver dans les exercices (Gras, Couturier, Guillet, & Spagnolo, 2005) (Card(M1)=66). Celui de la deuxième par sa rareté, elle reflète le choix de ce manuel de ne recourir que très rarement aux outils TICE (Card (M2)=2).
- N1, nature fréquentiste de l'objet. Ne pas proposer ni des situations de manipulation réelle ni de simulations semble ne laisser que très peu de chance aux situations fréquentistes (Card (N1)=1).
- P2, le calcul implicite de la probabilité. Le calcul d'une probabilité sollicité indirectement (Card (P2)=2) est une caractéristique de quelques exercices qui demandent de déterminer la valeur de la espérance mathématique, ils sont aussi très rare.
- D1, désigne les exercices dont la probabilité devient un outil de prise de décision (Card (D1)=0).

Arbre cohésitif. Manuel Nathan S

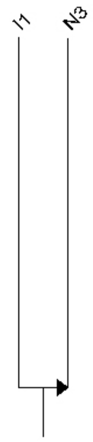


Graphique 26

Groupe H

La seule règle de cette classe (I1⇒N3) a été déjà décrite lors du groupe fréquentiste D, elle résume la relation conditionnelle *si il y a une interprétation de la probabilité, l'objet est ambigu*.

Ni Nathan S ni Belin S ne proposent aux élèves de produire leurs propres données (manipulation réelle), la voie alternative d'une simulation (M2) est plus fortement entreprise par Belin S (Card(M2)=15) que par Nathan S (Card(M2)=2) cela explique la disparition de M2 dans l'arbre cohésitif de Nathan S et présente par rapport à l'arbre de Belin S.



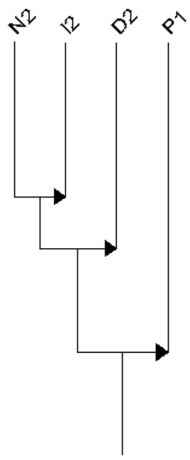
Cod	**Var**	**Mod**	**Card**
I1	**Inter P**	**Oui**	**6**
N3	**Nat A**	**Amb**	**22**

Tableau 36

En fin, cette classe nous renseigne plus pour cette variable absente (M2) que pour la construction de métarègles. D'ailleurs, même si le groupe optimal d'individus typiques pour cet algorithme donne un cardinal plus ample (5 individus) que le graphe implicatif (1 individu), les variables secondaires rendent compte des mêmes caractéristiques des exercices longs (L2 t.a.r. 0,154) avec un accent mis sur le calcul de l'espérance mathématique (T2 t.a.r. 0,0225).

Groupe I

Ce groupe est associé à la notion bayésienne de la probabilité, toutes ses variables sont incluses dans le groupe B de l'arbre de similarités et dans les groupes E et F du graphe implicatif. La première règle indique que *si l'objet est de nature générique alors l'interprétation de son calcul n'aura pas lieu,* relation qui amène à ne pas attendre une prise de décision concernant la probabilité; et finalement à demander explicitement le calcul de la probabilité.



Cod	**Var**	**Mod**	**Card**
I2	**Inter P**	**Non**	**56**
N2	**Nat A**	**Génér**	**39**
D2	**Décis**	**Non**	**62**
P1	**Calc P**	**Expl**	**64**

Tableau 37

Cette configuration de métarègles se trouve déjà dans son homonyme du manuel Belin S. Ces manuels semblent donc avoir entrepris la même voie pour traiter la notion bayésienne : des épreuves génériques en nombre important dont l'interprétation ne sera pas l'objet d'intérêt. Des exercices dont on arrive au calcul de la probabilité par une voie directe, sollicités explicitement et sans reprise sémantique, ni d'utilité décisionnelle non plus.

En proportion, ces exercices sont plus fréquents chez Nathan S que chez Belin S, le premier de ces manuels réduit la proportion d'exercices sans contextes et fréquentiste, laissant plus de place aux épreuves génériques. Néanmoins, ces caractéristiques contextuelles si récurrentes n'ont pas finalement de relevance, elles n'interviennent pas dans les questions posées et ne semblent être utilisées que pour leur facilité à créer des conditions d'incertitude nécessaires au concept de probabilité.

En effet, les épreuves génériques ont une certaine utilité didactique, ce sont des contextes relativement précis à évoquer et en même temps faciles à moduler, ceci permettant le traitement d'une gamme importante de tâches calculatoires à un coût relativement réduit.

Tel que ce qui arrive pour le manuel Belin S, les exercices de Nathan S se focalisent en général sur des stratégies combinatoires pour déterminer les cardinaux des cas favorables et possibles ou l'application des axiomes de Kolmogorov.

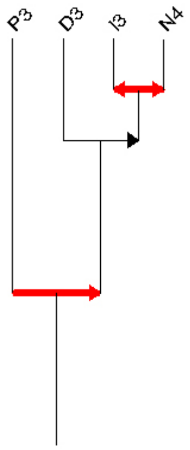
D'ailleurs, pour tous ces contextes, les événements sont de type unidimensionnel, en d'autres termes pour tout événement il n'y a qu'un cardinal possible, un seul ensemble de référence. De cette manière, admettant le principe d'équiprobabilité, on n'a qu'une seule possibilité pour évaluer la probabilité. Il n'y a aucun exercice des caractéristiques du cas de Camille, dont ses probabilités de réussite à un examen sont en fonction de l'ensemble de référence choisi.

Ces exercices, toujours d'événements unidimensionnels et d'application immédiate du principe d'équipossibilité laisseraient entendre que la probabilité d'une épreuve générique est quelque chose d'objectif, unique et, d'un point de vue didactique, plausible d'une correction en termes de "réponse correcte" ou "réponse incorrecte". Certains verbes utilisés (prouvez, déduisez, déterminez, etc.) et les conjugaisons d'autres (est, a, etc.) semblent contribuer à créer une sorte d'atmosphère déterministe, dont l'incertitude ne fait que s'adresser à l'issue de l'événement.

L'évocation du contexte (au lieu de la manipulation réelle) semble jouer un rôle non négligeable dans le contournement de la notion bayésienne. Elle facilite la relativisation de la nature de l'objet sur lequel porte le calcul de probabilité, pouvant-il passer d'épreuve générique vers une série infinie par l'intervention de l'enseignant. En effet, les contextes de ces exercices laissent toujours une place à une éventuelle adéquation à l'approche fréquentiste.

Groupe J

Ce groupe rassemble les exercices sans contextes ou qui ne concernent pas la notion de probabilité. Ces mêmes variables se sont trouvées déjà ensemble dans d'autres groupes : Groupe C (Analyse de similarités), Groupe G (Analyse implicative) et groupes homonymes des analyses du manuel Belin S.



Cod	Var	Mod	Card
P3	**Calc P**	**Non C**	**7**
D3	**Décis**	**Non C**	**11**
I3	**Inter P**	**Non C**	**11**
N4	**Nat A**	**Non C**	**11**

Tableau 38

Cette classe confirme les conclusions tirées lors des analyses précédentes, son groupe optimal d'exercices typiques inclut l'exercice 10, transcrit déjà lors de l'interprétation de la classe C de l'analyse de similarités. Ce groupe complémente donc la caractérisation des exercices sans contextes ou sans traitement de la probabilité. Les mêmes variables secondaires typiques nous confirment cette hypothèse, il s'agit d'exercices courts (T1 t.a.r. 0,166) posant peu de questions (Q1 t.a.r. 0,0978) avec une attention particulière sur le registre symbolique (R3 t.a.r. 0,179) dont les probabilités sont présentées directement (H2 t.a.r. 0,0121) du au manque d'un contexte permettant de les déterminer autrement.

Enfin, les trois analyses effectuées dans le manuel Nathan S nous ont permis d'identifier trois groupes d'exercices, l'un fréquentiste, l'autre bayésien et un dernier sans contexte. Les relations entre les variables de ces groupes nous informent de la manière de résoudre le conflit entre la dualité de la probabilité et l'intention des programmes de ne reconnaître que la fréquentiste.

La solution proposée par ce manuel, similaire à celle de Belin S, consiste à gérer le statut institutionnel de chacune des notions par les questions posées. Les exercices de nature fréquentiste posent des questions autour de la notion en jeu ; ceux de nature bayésienne n'interrogent pas sur les aspects sémantiques. A la place de ces derniers, l'attention se centre sur les calculatoires. Nous résumons les caractéristiques des trois types d'exercices pour le manuel Nathan S :

Groupe fréquentiste

Tel qu'il est observé dans l'autre manuel de la même filière, Nathan S présente un groupe d'exercices fréquentistes. Celui-ci se manifeste dans les trois analyses effectuées : Groupe A (Analyse de similarités), Groupes D–F (Analyse implicative) et Groupe H (Analyse cohésitive).

Le manuel Nathan S combine deux choix qui pénalisent les exercices fréquentistes, d'une part l'absence de contextes empêche d'étudier matériellement la stabilisation de fréquences, et d'autre part, la voie de sa substitution par une simulation est très rarement entreprise. De cette manière les exercices dont la nature de l'objet est la série infinie n'ont pas pratiquement des possibilités se manifester.

Toutefois, cette faible quantité d'exercices fréquentistes a tendance à ressembler à ceux du manuel Belin S. En effet, ils sont les plus longs des trois catégories considérées, suivis par les bayésiens et finalement les « sans contexte ». D'ailleurs, quelques exercices fondent les deux notions par l'utilisation du principe fréquentiste dans l'évaluation d'une épreuve générique (nature ambiguë).

Groupe bayésien

Cette classe apparaît elle aussi sur les trois outils statistiques utilisés : Groupe B (Analyse de similarités), Groupe E-F (Analyse implicative) et Groupe I (Analyse cohésitive).

Ce groupe a une présence plus forte en proportion en relation à son homonyme de Belin S, dû à la rétraction de ces deux suppléments (groupes fréquentiste et sans contexte). Néanmoins, le contexte de l'exercice n'a qu'une fonction secondaire de décor pour encadrer des tâches calculatoires sur la probabilité. Les questions de cette masse importante d'exercices ne s'occupent pas des aspects sémantiques associés aux calculs de probabilité effectués, ces exercices deviennent finalement du type technique dont le contexte incertain répond à un requis de la notion de probabilité, en particulier pour la différencier d'une simple proportion.

Paradoxalement, ces exercices introduisent un lexique déterministe. En effet, en présentant des événements unidimensionnels, et en s'appuyant sur le principe d'équipossibilité, ils laissent entendre une certaine unicité et déductibilité dans les procédés d'évaluation d'une probabilité, ainsi il est fréquent de trouver des termes tels que prouvez, déterminer, déduisez, etc. dans les énoncés de ces exercices bayésiens.

Groupe sans contexte

Ce groupe aussi a été présent lors des trois analyses effectués avec le logiciel CHIC : Groupe C (Analyse de similarités), Groupe G (Analyse implicatif) et Groupe J (Analyse cohésitive).

Par leur caractéristiques, ces exercices sont étrangers à toute tentative d'interprétation de la probabilité, ils se dispensent parfois du contexte ou des conditions d'incertitude ou même de la caractérisation de l'événement, la tâche principale se focalisant sur des aspects mathématiques dont leur résolution ne nécessite que des propriétés mathématiques.

Bref, des caractéristiques communes semblent se dessiner à partir de l'analyse des deux manuels de la filière Scientifique, la dualité de la probabilité est gérée par les questions posées, autant que de variables didactiques, pointent sur la notion fréquentiste lorsque le contexte ainsi l'admet ou sur des tâches calculatoires pour les contextes bayésiens.

Nous analyserons un troisième manuel, cette fois-ci de la filière Economie et sciences sociales.

2.7 Manuel Bréal ES

Ce manuel de 287 pages consacre 31 exercices au chapitre Probabilité, ils sont repartis en deux sections, la première appelée *Maîtriser le cours*, la deuxième *Savoir-faire fondamentaux*. Le tableau ci-dessous montre la distribution des exercices par modalité.

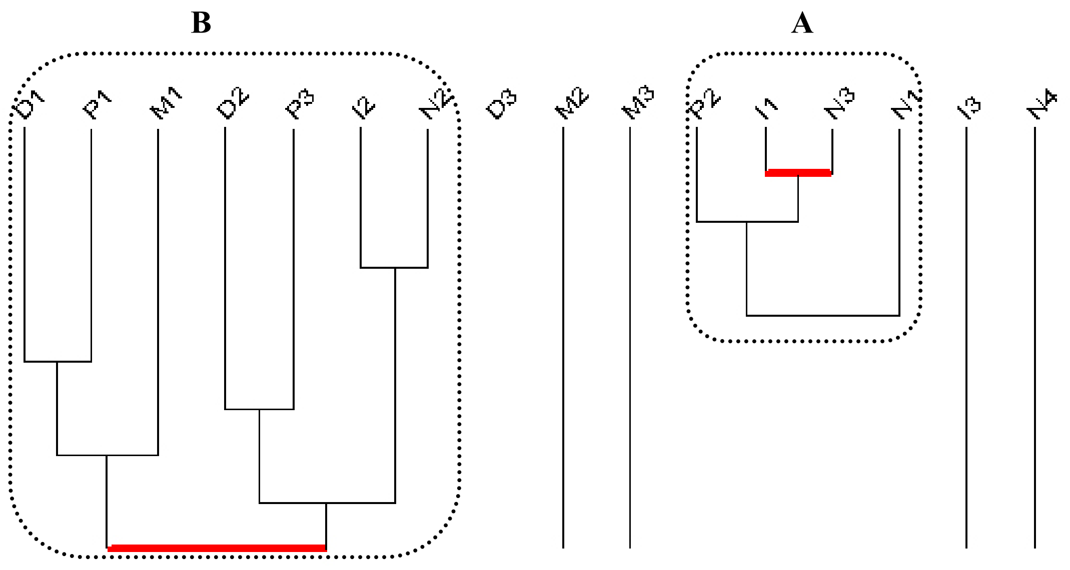
Distribution du nombre d'exercices du manuel Bréal ES par modalité

Lignes		Questions			Représentations				Tâches				Contexte				
L1	L2	Q1	Q2	Q3	R1	R2	R3	R4	T1	T2	T3	T4	C1	C2	C3	C4	C5
20	11	5	13	13	6	0	2	23	29	1	0	1	16	13	0	2	0

Décision			Hypothèses				Manipulation			Calcul P			Interprétat P			Nature A			
D1	D2	D3	H1	H2	H3	H4	M1	M2	M3	P1	P2	P3	I1	I2	I3	N1	N2	N3	N4
2	29	0	22	5	3	1	31	0	0	29	1	1	9	22	0	1	22	8	0

Arbre de similarités

Arbre de similarités. Manuel Bréal ES



Graphique 28

Les exercices des deux premiers manuels analysés sont regroupés toujours autour de trois classes, une fréquentiste, autre bayésienne et une dernière, d'exercices sans contexte. En observant l'arbre de similarités du manuel Bréal ES nous découvrons la dilution de la dernière de ces classes. Cette absence s'explique par le choix de ce manuel de ne pas proposer d'exercices sans contexte (Card(C5)=0), ainsi plusieurs de variables dérivées de C5 ne se sont pas manifestées (Card(N4)=0, Card(M3)=0 et (Card(P3)=0). Il reste donc deux classes, la première fréquentiste (Groupe A) qui commence à se constituer à partir du couple de variables (I1, N3) et la deuxième bayésienne (Groupe B) qui le fait par (I2, N2).

Groupe A

Cette classe, homonyme à celles des analyses respectives des manuels de la filière Scientifique, se caractérise par des exercices que décrivent des contextes dont la nature de l'objet devient ambiguë (N3) avec un intérêt pour l'interprétation de la probabilité calculée (I1). Ce couple de variables constituant le noyau de la classe, d'un point de vue tant sémantique que statistique. En effet, ce premier rassemblement nous informe que ces caractéristiques se corrèlent fortement dans ces exercices formant ainsi le premier couple de variables ressemblées, on doit s'attendre donc à ce que lorsque un exercice s'intéresse à une interprétation de la probabilité, cet objet soit quelque part ambigu et vice-versa. Néanmoins, malgré cette forte relation symétrique, cette sous-classe reste faible en nombre d'exercices (Card($I1 \cap N3$ =5).

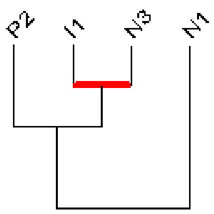


Tableau 39

Pour cette classe nous transcrivons un des cinq exercices de son groupe optimal d'individus typiques (Exercice 27, Bréal ES, page 102) :

> *« Un couple envisage d'avoir au maximum quatre enfants, mais décide –on ne sait pas pourquoi- de s'arrêter à la naissance de la première fille.* ***On s'intéresse au nombre d'enfants de cette famille*** (en gras dans l'original).
>
> *1. Quelle est la probabilité que cette famille ait : un enfant ? deux enfants ? trois enfants ? quatre enfants ? On explicitera les hypothèses de la modélisation faite pour conduire les calculs.*

2. *Un démographe a cherché dans le registre d'état-civil d'une ville les familles du type F, GF, GGGF ou GGGG. Il a obtenu les renseignements suivants :*

Nombre d'enfants	**Nombre de familles**
1	**8588**
2	**4336**
3	**2222**
4	**2417**

Ces données sont-elles en bonne adéquation avec les calculs de la question précédente ? Peut-on conserver ce modèle ?
3. *Les données statistiques générales montrent que la fréquence de naissance d'un garçon est légèrement supérieure à celle de la naissance d'une fille. Reprendre l'exercice avec l'hypothèse p(G)=0 ,51 et P(F)=0,49.*
Les résultats obtenus par les calculs conduits dans ce modèle sont-ils en bonne adéquation avec les données du tableau précèdent ? ».

La relation entre la fréquence d'apparition et le calcul de probabilité (I3) s'établit intentionnellement à plusieurs reprises, néanmoins, la fréquence y a des rôles différents. En effet, l'exercice commence en s'intéressant au nombre d'enfants d'une famille qui, par le manque d'information ou de signe distinctif, est perçue comme un couple parmi d'autres, en autres termes un couple générique. L'évaluation de la probabilité se réalise donc par un ensemble de référence qui pourrait être fini (cas favorables sur cas possibles) ou infini (fréquence d'apparition), cette dernière possibilité est la retenu dans la deuxième partie de l'exercice.

Il y a donc une confluence de deux objets, le premier est une épreuve générique, le deuxième la série infinie. Cet exercice est, et par son récit, un des rares à le montrer si clairement. La différence d'approches est même accentuée par des acteurs différents, d'un côté le couple et son intérêt pour le genre de son prochain bébé et de l'autre le démographe qui s'intéresse à la fréquence d'apparition, son objet est la population. L'exercice, en s'appuyant sur la naturalité d'application de ce principe, laisse sans dévoiler ce procédé.

L'évocation, encore une fois, semble faciliter cette sorte de relais entre objets, en centrant l'attention sur un objet, l'abandonnant ensuite, allant à s'intéresser à un autre pour enfin revenir sur le premier, la liaison se fait par la naturalité de l'application du principe fréquentiste. Cette liaison se fait par la gestion en temps et en forme de la présentation des données de l'exercice.

L'évocation n'a pas seulement la fonction de fusion d'objets, elle a la fonction complémentaire de blocage d'autres démarches de raisonnements, un couple est présenté, mais son absence fait abandoner toute tentative de recherche de données particulières, il devient ainsi générique, puis l'attention pivote vers les naissances en la population, son absence empêche entre autres, toute question concernant la légitimité de l'appartenance du couple en question à cette population, et finalement l'attention revient sur le couple. Si à ce contexte restreint en information on additionne a) la naturalité du principe fréquentiste, b)un contrat didactique qui force le lien entre les données, et c) des questions qui n'interrogent pas au niveau sémantique, alors nous ne devrions pas nous étonner que ces problèmes soient probablement résolus sans difficultés majeures.

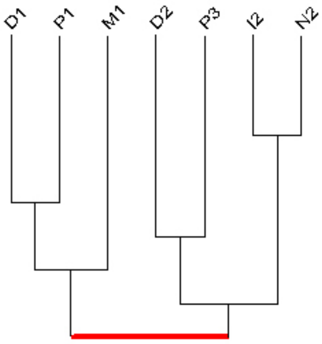
Dans nos expérimentations, nous nous intéresserons aux contextes réels et évoqués en tant que variables didactiques qui gèrent des possibles réactions et questions des élèves, pour l'instant nous revenons sur la classe fréquentiste de ce manuel de la filière Economie et sciences sociales.

Pour ce qui concerne sa représentativité dans son groupe, cet exercice l'est par l'utilisation de tableaux aux fins organisationnelles (R1 t.a.r. 0,0572), par sa longueur (L2 t.a.r. 0,256) et pour poser un nombre non négligeable de question (Q2 t.a.r. 0,351). A cette classe se joignent, par un seul exercice, les variables P2 (Calcul implicite de la probabilité) et N1 (la série comme nature de l'objet à probabiliser).

Tel que les groupes fréquentistes des manuels de la filière Scientifique, ceux-ci sont aussi très peu nombreux, nous nous intéresserons à un autre groupe qui se dessine de l'arbre de similarités du manuel Bréal ES.

Groupe B

Cette classe s'est vue élargie en termes de variables, par rapport à ses homonymes des manuels précédents, principalement par la présence de la variable D1 (prise de décision). En effet, deux exercices mobilisent les principes d'un test d'hypothèse sur la proportion d'une population par celle d'un échantillon donnée dans l'énoncé.



Cod	Var	Mod	Card
D1	Décis	Oui	2
P1	Calc P	Expl	29
M1	Manip	Evoq	31
D2	Décis	Non	29
P3	Calc P	Non	1
I2	Inter P	Non	22
N2	Nat A	Génér	22

Tableau 40

Certains regroupements par paires de variables sont assez asymétriques en cardinalité ((D1, P1) et (D2, P3)), nous aurions souhaité présenter un exemple extrait d'un groupe optimal d'une sous-classe moins déséquilibrée ((I2, N2), (Card($I2 \cap D2$)=22), mais finalement, son groupe optimal coïncide avec celui de la classe complète, nous en transcrivons un exemple (exercice 4, Bréal ES, page 98) :

« Dans un village de vacances il y a 95 adultes (donc 60 femmes) et 45 enfants (dont 25 filles). Un enquêteur interroge au hasard une personne de ce groupe. Quelle est la probabilité que la personne interrogée :

a) soit adulte

b) soit de sexe masculin

c) soit adulte ou de sexe masculin

d) ne soit adulte ni de sexe masculin »

Cet exercice constitue un exemple typique d'épreuve générique (N2) dont l'évaluation de la probabilité, explicitement sollicitée (P1), se réalisera par le principe d'équipossibilité.

Tel qu'il arrive pour les groupes B des autres manuels, l'énoncé n'interroge pas sur la sémantique de l'objet calculé (I2), ni de la place la probabilité comme outil de prise de décision (D2).

Cet exemple illustre les caractéristiques des épreuves génériques déjà repérées lors des analyses des manuels précédents, en particulier celles qui concernent les contextes concis et au même temps facilement modulables à des fins calculatoires. Néanmoins cette classe se différencie des bayésiennes des autres manuels par un certain nombre de particularités : les exercices ont tendance à poser un nombre non négligeable de questions (Q3 t.a.r. 0,194) et à interroger sur les valeurs des paramètres de la population à partir des outils calculés (H3 t.a.r. 0,23).

Comme nous l'avons signalé précédemment, l'absence d'exercices sans contexte (C5) isole un ensemble de variables (N4, M3, D3 et I3), contrairement aux autres manuels qui rencontraient un élément de liaison.

Ce groupe bayésien, d'épreuves génériques, reflète par quelques variables (D1 et P3) la singularité de ce manuel de la maison Bréal par rapport à ceux de la filière Scientifique. Les autres outils statistiques du logiciel CHIC nous permettrons de complémenter l'information obtenue par l'analyse de similarités, nous continuons avec le graphique implicatif qui nous renseigne sur les relations non symétriques entre variables.

Graphe implicatif

L'analyse implicative nous confirme l'impression que nous a donné ce manuel lors des analyses *a priori* de ses exercices, ils nous ont paru quelque part différents et plus variés. Sa diversité provient d'une approche relativement singulière qui se manifeste entre autres par le faible nombre d'exercices proposés (31), qui y sont plus longs que chez les autres manuels et posent plus de questions ($\text{Card}(Q2 \cup Q3)=26$). D'ailleurs, certaines variables n'avaient été observées que de manière isolée, les décisions à entreprendre ($\text{Card}(D1)=2$), le manque de simulations ($\text{Card}(M2)=0$) ou même l'absence d'exercices sans contexte ($\text{Card}(C5)=0$) sont les signes d'une approche singulière.

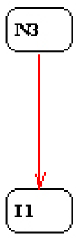
La combinaison du faible nombre d'exercices et leur grande diversité fait de la régularité une caractéristique difficile à trouver. C'est ainsi qu'il nous a fallu descendre le seuil à 0,65 pour qu'une première règle émerge de ce manuel ($N3 \Rightarrow I1$). Le couple de variables de cette implication se trouve dans tous les groupes D (fréquentiste) des deux manuels de la filière S mais inversée ($I1 \Rightarrow N3$), nous décrirons cette relation conditionnelle.

Groupe D

La règle N3⇒I1 représente une proportion non négligeable d'exercices ($\text{Card}_{N3}(I1) = 5/8$). D'ailleurs, la ressemblance entre ces deux variables est la plus forte de l'arbre de similarités (Tableau 41) ($\text{Card}_{I1}(N3) = 5/9$). Ainsi ce manuel fait devenir condition de quasi suffisance ce qui était dans les autres de nécessité.

Cette règle (*si la nature de l'objet est ambiguë, alors il y a une interprétation de la probabilité)* n'est pas, en termes sémantiques, si significative que sa réciproque. Toutefois elle nous renseigne sur l'ample utilisation du principe fréquentiste dans des situations qui alternent entre épreuve générique et série infinie. L'exemple transcrit pour le groupe A, qui se trouve d'ailleurs dans l'intersection de ces deux variables, nous a montré comment un énoncé réalise, et ce pour un même calcul de probabilité, un relais d'objets (des générique vers la série et vice versa).

Graphe implicatif (seuil 0,65). Manuel Bréal ES



Cod	Var	Mod	Card
I1	Inter P	Oui	9
N3	Nat A	Amb	8

Tableau 41

D'ailleurs, lors des analyses *a priori* nous avons eu des difficultés à repérer quelques modalités, en particulier N3 (Nature de l'objet ambiguë). Son identification n'est pas si facile à cerner chez Bréal ES, ses énoncés sont beaucoup moins structurés que ceux des manuels précédents, les événements sont rarement définis de manière à que nos variables puissent prendre aisément leurs valeur.

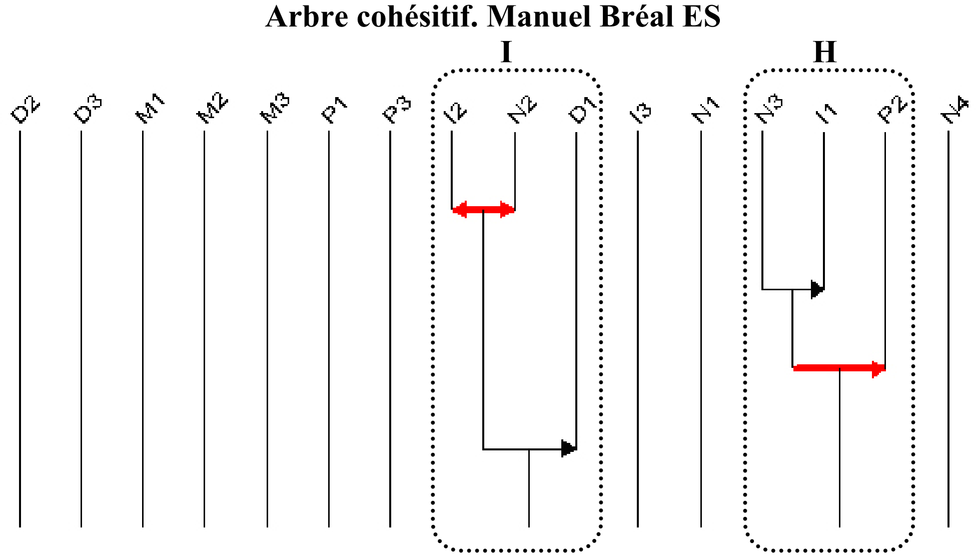
Pour ce manuel en particulier, la modalité N3 représente parfois notre impossibilité à préciser l'objet en question, tandis que pour les autres, la situation étant toute autre, nous avons détecté, sans majeure difficulté, les deux objets et leur alternance respective. Pour le manuel Bréal la situation est différente, et bien qu'il puisse y avoir un « effet de filière », de relaxation dans les énoncés, il est probable aussi que nos variables s'avèrent insuffisantes face à cette structure d'énoncés si particulière.

A la diversité signalée précédemment, nous ajoutons donc une certaine inadéquation de nos variables devant ces exercices, cette combinaison semblant contribuer au faible nombre de règles repérées pour le manuel Bréal ES. Nous continuons pour l'arbre cohésitif à la recherche de métarègles.

Arbre cohésitif

Pour ce manuel, l'arbre cohésitif regroupe peu de métarègles. Par leurs sémantiques ces classes ressemblent respectivement au groupe fréquentiste (Groupe H) et au bayésien (Groupe I) laissant un nombre très important de variables isolées. Nous commençons par le groupe fréquentiste.

Arbre cohésitif. Manuel Bréal ES



Graphique 30

Groupe H

Ce groupe réunit des exercices dont l'interprétation de la probabilité se fait sur des situations ambiguës qui à leur tour demandent de la calculer de manière implicite, la première quasi-implication ayant déjà été repérée lors de l'arbre implicatif (N3⇒I1).

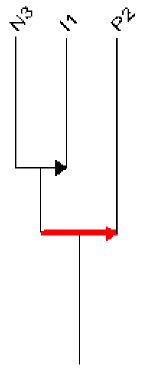


Tableau 42

Cod	Var	Mod	Card
P2	Calc P	Impl	1
I1	Inter P	Oui	9
N3	Nat A	Amb	8

L'exercice que nous avons transcrit pour le groupe fréquentiste A de l'arbre de similarités ne se trouve pas dans le groupe optimal d'individus typiques de cette classe, nous en profitons alors pour transcrire un autre exemple (Exercice 13, Bréal ES, page 99) :

« Une pièce lancée dix fois a présenté dix fois le côté face. A la question : « quelle est la probabilité que cette pièce présente « face » si on la lance une fois de plus ? », Pierre répond « 0,5 bien sûr !». Expliquer la réponse de Pierre et la commenter. »

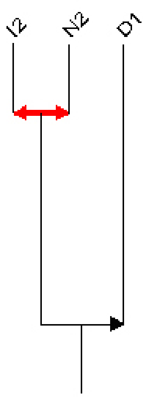
Tout d'abord, cet exercice n'est pas celui qui demande de calculer la probabilité de manière implicite (P2). Néanmoins il est un de ceux qui évoquent des objets de nature différente (N3). D'une part il y a l'objet fréquentiste qui joindra la discussion par la douteuse représentativité d'un échantillon si faible, et d'autre part le prochain lancer, objet qui par sa nature est bayésien.

Bien que cet exercice présente l'originalité des réponses ouvertes, il répète le modèle des épreuves génériques dont la fréquence est utilisée pour évaluer une mesure de certitude portée sur un événement singulier.

Finalement, pour cette classe, une seule différence est à repérer par rapport à ses homonymes des autres manuels : elle concerne la longueur des énoncés qui ont tendance a être plus courts qu'ailleurs (Q1 t.a.r 0,128). Nous analysons l'autre groupe de l'arbre cohésitif du manuel Bréal ES.

Groupe I

La relation de presque équivalence entre I2 et N2 se distingue dans cette classe (Card$_{N2}(I2)$= Card$_{I2}(N2)$=19/22). Cette relation de forte corrélation nous informe d'exercices de nature générique dont la probabilité n'est pas interprétée et vice-versa. Nous retrouvons encore une fois dans un groupe bayésien l'association de ces deux variables.



Cod	**Var**	**Mod**	**Card**
I2	**Inter P**	**Non**	**22**
N2	**Nat A**	**Génér**	**22**
D1	**Décis**	**Oui**	**2**

Tableau 43

Cette double quasi-implication rejoint faiblement D1 (des exercices dont la probabilité calculée intervienne dans une prise de décision). Du groupe optimal d'éléments typiques du groupe I (28 exercices ; 90% de ceux du chapitre probabilité), nous transcrivons un exercice avec prise de décision (Exercice 14, Bréal 11, page 99) :

« Lors de 100 lancers d'une pièce, « pile » a été obtenu 57 fois.

1. *Peut-on modéliser le fonctionnement de cette pièce en la supposant équilibrée ? Sinon, proposer une autre modélisation, en expliquant les raisons.*
2. *Calculer dans le modèle choisi la probabilité d'obtenir deux fois « face» lors de deux lancers de la pièce.*

Lorsque nous nous intéressons à la probabilité en tant qu'élément de décision, cela inclut le sens usuel attribué dans la théorie de la décision. Mais nous ne nous sommes pas restreint à cette seule théorie, le sens que nous lui avons assigné est plus large et concerne toute action à entreprendre dont la probabilité participe à l'argument qui fonde l'action en question. Pour cet énoncé, l'action pourrait se traduire comme celle d'accepter un modèle ou un autre, la fréquence d'apparition interviendrait alors pour argumenter le choix du modèle.

L'objet sur lequel porte le calcul de la probabilité (Question 2) est une épreuve générique : son calcul n'étant pas soumis à une interprétation, l'exercice finit immédiatement après l'avoir sollicité. Pour ce qui concerne donc l'interprétation de la probabilité calculée et s'agissant d'une épreuve générique, ce problème est similaire à ceux des autres manuels dont le calcul n'est pas interprété, l'originalité se trouvant dans le possible débat autour du choix du modèle.

D'ailleurs, cette classe ne possède pas de variables secondaires typiques au-dessus d'un risque de 0,5, ce qui illustre la diversité des énoncés de ce manuel.

Enfin, les trois analyses effectuées sur le manuel Bréal ES nous renseignent d'une part de la singularité de l'approche de ce manuel (exercices ouverts, prises de décisions, énoncés tous contextualisés, etc.), et d'autre part sur le recours à la même stratégie que les autres manuels pour traiter la dualité de la probabilité, autrement dit, un contournement de la notion bayésienne par le type de questions posées. Par la suite, nous décrirons les groupes fréquentiste et bayésien de ce manuel.

Groupe fréquentiste

Ce groupe se manifeste dans les trois analyses effectuées : Groupe A (Analyse de similarités), Groupes D (Analyse implicative) et Groupe H (Analyse cohésitive).

Tel que les autres groupes fréquentistes repérés chez les manuels de la filière Scientifique, celui-ci est aussi faiblement représenté, l'intersection des deux variables centrales de ce groupe (I1 : interprétation de la probabilité et N3 : objets de nature ambiguë) nous renseigne que pas plus de 16 % des exercices répondent à cette caractéristique conjointe.

Toutes les contextes de ce manuel sont évoqués, de plus, les auteurs ont choisi de ne pas introduire de simulations, cette combinaison compliquerait la création de contextes purement fréquentistes. Ainsi la variable N1 (Card(N1)=1) disparaît de cette classe et il ne reste que N3 (Card(N3)=8).

La modalité N3 qui dans les autres manuels caractérise les exercices de nature ambiguë, chez Bréal ES acquiert ici une autre signification, elle réunit aussi ceux pour lesquels il est difficile de préciser la nature de l'objet sur lequel porte la probabilité. En fait, un seul exercice répond strictement à la modalité N1 (série infinie), les autres se partagent entre ceux qui convoquent deux objets de nature différente et ceux pour lesquels il devient difficile de déterminer l'objet en question.

Un autre groupe, bayésien cette fois-ci, s'observe au travers des analyses du manuel Bréal ES, nous en décrivons brièvement ses caractéristiques.

Groupe bayésien

Cette classe se manifeste sur deux des trois outils statistiques utilisés : Groupe B (Analyse de similarités) et Groupe I (Analyse cohésitive). L'analyse de similarités, moins restrictive que la dernière, réunit 7 variables tandis que la cohésitive juste 3.

Dans ces deux groupes, le rassemblement commence par le couple (I2, N2), dans l'arbre cohésitif s'observent des quasi implications dans les deux sens (I2$\Rightarrow$ N2 et N2$\Rightarrow$ I2) ce qui renvoie à la notion de similarité entre I2 et N2. L'intersection de ces deux variables représente presque 71 % des exercices du manuel. C'est ainsi qu'un grand nombre d'exercices sont strictement de nature générique (N2) et ne posent pas de questions sur les calculs réalisés (I2).

Le groupe bayésien, comme dans les autres manuels, se différencie en partie du groupe fréquentiste par la non évocation de la fréquence d'apparition au moment de l'évaluation de la probabilité. En d'autres termes, dans tous les groupes fréquentistes, nous avons trouvés des exercices avec concurrence d'objets. Pour ces derniers, la demande du calcul de la probabilité se fait sur une épreuve générique et se valide par le principe fréquentiste, généralement ce principe apparaît *a posteriori* du calcul réalisé par le principe d'équipossibilité.

Cela veut dire qu'à l'intérieur des groupes fréquentistes il y a des questions de nature bayésienne. La différence entre les questions bayésiennes du groupe fréquentiste et celles du groupe bayésien est dans le critère d'évaluation utilisé : pour le premier il se mobilise le principe fréquentiste et pour le deuxième celui d'équipossibilité.

Autrement dit, rares sont les exercices de nature strictement fréquentiste. La plupart sont des objets du type épreuve générique, les uns évalués par le principe d'équipossibilité, les autres par le principe fréquentiste.

Ces deux manières d'aborder les épreuves génériques ont en commun la non reconnaissance de la notion bayésienne. Pour ceux qui évoquent la fréquence, le vrai rôle de cette interprétation n'est pas mis en évidence, rendant ainsi l'objet flou et prêt à une redéfinition fréquentiste si les circonstances le requièrent.

Nous analyserons le dernier manuel, lui aussi de la filière Economie et sciences sociales.

2.8 Manuel Didier ES

Ce manuel de 222 pages distribue ses 65 exercices du chapitre Probabilité en trois sections, *Travail autonome* (Exercices 1 à 5), *Pour s'entraîner* (Exercices 6 à 50) et *Pour chercher* (Exercices 51 à 65). Le tableau ci-dessous représente la répartition des exercices selon les différentes variables.

Distribution du nombre d'exercices du manuel Didier ES par modalité

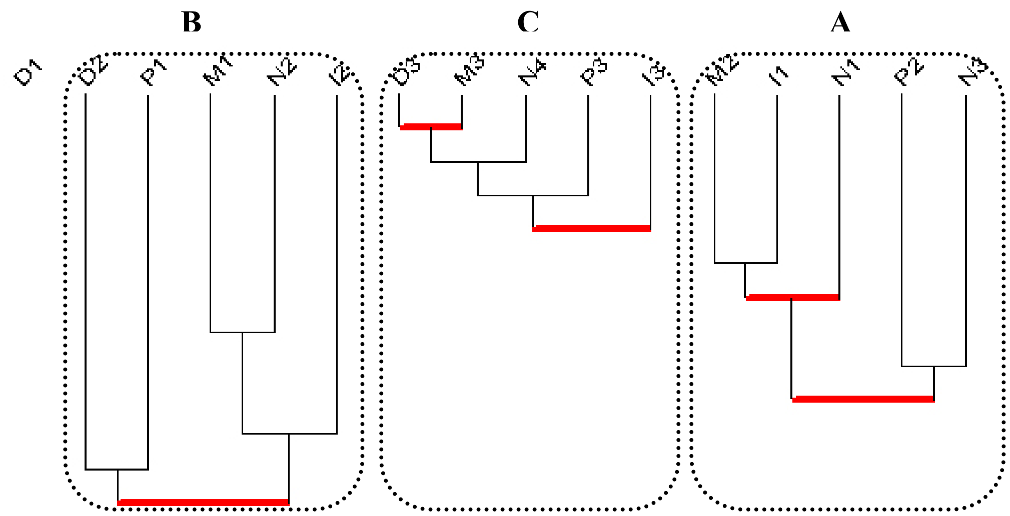
Lignes		Questions			Représentations				Tâches				Contexte				
L1	L2	Q1	Q2	Q3	R1	R2	R3	R4	T1	T2	T3	T4	C1	C2	C3	C4	C5
43	22	21	17	27	29	6	5	25	47	0	17	1	36	5	2	6	10

Décision			Hypothèses				Manipulation			Calcul P			Interprétat P			Nature A			
D1	D2	D3	H1	H2	H3	H4	M 1	M 2	M 3	P1	P2	P3	I1	I2	I3	N1	N2	N3	N4
0	55	10	50	8	0	7	41	11	13	49	6	10	12	45	8	8	36	9	12

Le manuel Didier de la filière ES ressemble plus à ceux de la filière S qu'à Belin ES, au moins par le nombre d'exercices. Nous commençons l'analyse de ce manuel par l'arbre de similarités. Sauf D1 (Prise de décision) toutes les variables font partie d'un des groupes de l'arbre, cela nous informe que ces exercices sont plus semblables entre eux que ceux proposés par le manuel Bréal ES.

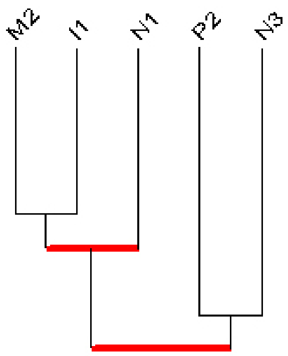
Arbre de similarité

Arbre de similarités. Manuel Didier ES



Groupe A

Ce manuel, et à la différence de Breal ES, introduit l'utilisation des TICE dans le chapitre Probabilité. De cette manière les objets de nature strictement fréquentiste (N1) réapparaissent dans l'arbre des similarités. En effet, rappelons le choix des auteurs du manuel précédent à ne traiter que très rarement l'utilisation de l'ordinateur ; ce choix, combiné à celui d'éviter les contextes réels a rendu difficile la construction de situations fréquentistes. Le manuel Didier ES ouvre la possibilité aux simulations, ce qui lui permet de créer des contextes où la nature de l'objet est la série (N1).



Cod	Var	Mod	Card
M2	Manip	Simul	11
I1	Inter P	Oui	12
N3	Nat A	Amb	9
P2	Calc P	Implic	6
N1	Nat A	Série	8

Tableau 44

La classe fréquentiste de cet arbre commence à se constituer par les variables N1 (Nature de l'objet : série) et M2 (simulations). Cette association nous parle du choix de ce manuel, de substituer la production de données réelles par celles de l'ordinateur.

Nous n'avons pas trouvé, ici non plus, des exercices proposant la manipulation réelle et sa simulation. En même temps tous les exercices, et cela concerne tous les manuels, présentent directement la situation simulée. Il reste à savoir si ces contextes simulés sont facilement transposables pour les élèves vers des contextes réels et en plus si des tels contextes sont si mobilisateurs de débats autour de questions déterministes

Cette classe fréquentiste du manuel Didier ES termine par les variables P2 (Calcul implicite de la probabilité) et N3 (nature ambiguë de l'objet). Nous aurions souhaité transcrire un exercice du groupe optimal des individus typiques de la sous-classe ((M2, I1) N1), vu qu'il est plus compact que la classe complète (((M2, I1) N1), (P2,N3)), mais finalement les deux groupes optimaux sont identiques. Nous en transcrivons un, afin d'illustrer les variables rassemblées (Exercice 54, Manuel Didier ES, page 210) :

« On jette deux dés cubiques bien équilibrés. On cherche la fréquence de sortie d'un double, c'est-à-dire par exemple 2 fois le 1, 2 fois le 2, etc.

1. Réaliser une feuille de calcul du même type que celle ci-dessous :

L1C3 = =SI(x=y,1,0)

	1	2	3
1	1	3	0
2	2	3	0

Effectuer un grand nombre de simulations et estimer la fréquence recherchée.

2. Calculer la probabilité de sortie d'un double en s'aidant d'un tableau ».

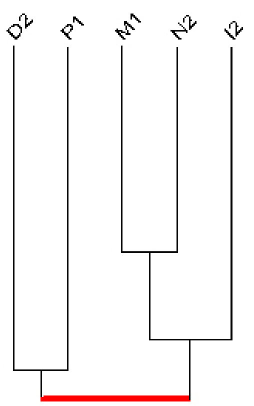
De contexte ludique, cet exercice ne propose pas d'évaluer la probabilité sur un événement singulier, il reste focalisé sur la série. C'est ainsi que la nature de l'objet est toujours N1. Des cinq variables de cette classe, cet exercice répond aux trois premières (M2, I1 et N1), celles qui en fait constituent le noyau central de la classe fréquentiste.

Les variables secondaires typiques de cette classe, nous renseignent sur des exercices relativement courts (L1 t.a.r. 0,365) dont les contextes sont plutôt des jeux de hasard (C1 t.a.r. 0,11), posant généralement entre trois et quatre questions (Q2 t.a.r. 0,0161) et avec une forte présence de tableaux dans les énoncés (T3 t.a.r. 0,00312).

Les 8 exercices de la variable N1 sont tous des contextes du type C1 (jeux de hasard), et sans aucune référence à des possibles décisions à entreprendre avec les calculs effectués. Tel qu'il a été observé auparavant, l'élève est invité à réaliser des calculs sans dépasser la position d'observateur de la pseudo réalité. Nous continuons l'analyse de l'arbre de similarités par le groupe B de ce manuel.

Groupe B

Ce groupe bayésien est significativement plus nombreux que le précédent. La correspondance entre modalités nous indique que ces exercices triplent ceux de la classe fréquentiste.



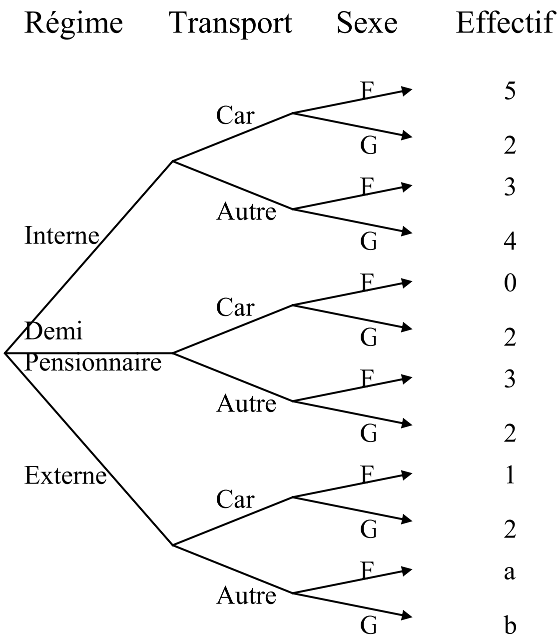
Cod	Var	Mod	Card
I2	**Inter P**	**Non**	**45**
N2	**Nat A**	**Génér**	**36**
D2	**Décis**	**Non**	**55**
M1	**Manip**	**Evoq**	**41**
P1	**Calc P**	**Expl**	**49**

Tableau 45

Ce groupe, le moins homogène des trois, se caractérise principalement par des exercices dont les objets sont de nature épreuve générique et comme toujours, évoqués. Les calculs de probabilités se sollicitent de manière explicite et ne sont pas interprétés, les variables de son homonyme du manuel Belin S.

Le cardinal du groupe optimal d'exercices typiques est assez nombreux, de ses 36 exercices nous en transcrirons un intitulé *« Les élèves sont fichés »* :

« Voici un arbre qui représente la répartition des 35 élèves d'une classe selon leur régime, leur mode de transport et leur sexe :



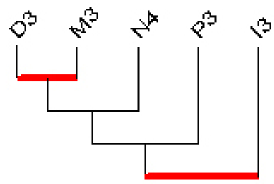
1. *Le nombre d'élèves des deux dernières branches de l'arbre a été oublié. Retrouver a et b sachant que dans la classe il y a 16 filles.*
2. *On choisit au hasard une fiche d'élève. Quelle est la probabilité d'obtenir :*
   a. *Une fille ?*
   b. *Un demi-pensionnaire ?*
   c. *Un élève prenant le car ?*
   d. *Une fille prenant le car et qui soit interne ?*
   e. *Un garçon externe ne prenant pas le car ?*

Cet exercice de type épreuve générique montre les avantages de la modularité à laquelle nous avons fait référence lorsque nous essayons d'expliquer l'ample présence de ces exercices dans ces manuels. Ils sont de grande utilité par la facilité à déplier un nombre considérable de tâches tout en posant des questions concises. C'est ainsi que pourrait se justifier la présence des exercices bayésiens. Ils ont pour objet d'entraîner les élèves à réaliser des calculs à partir du principe d'équipossibilité et des axiomes de Kolmogorov.

Enfin, cette classe se caractérise par plusieurs variables secondaires typiques, contrairement à la classe homonyme des manuels de la filière Scientifique, ces exercices sont longs (L2 t.a.r. 0,0755), posent un nombre important de question (Q3 t.a.r. 0,00108) et se ressemblent par un travail important sur des notions ensemblistes (T1 t.a.r. 0,00172, H1 t.a.r. 0,0188). En effet, les manuels de la filière Economie et sciences sociales ne proposent pas d'exercices faisant intervenir l'espérance mathématique, cet objet relativement complexe est exclusif aux manuels de la filière S ; les livres de texte de la filière ES circonscrivent les tâches à des éléments de dénombrement, de combinatoire et aux axiomes de Kolmogorov. Nous décrivons le troisième groupe de l'analyse de similarités.

Groupe C

Ce groupe, le plus compact des trois, réunit des exercices caractérisés par une absence de contexte (Card(C5)=10). Tel que nous l'avons expliqué pour les autres manuels, à partir de cette modalité s'enchaînent d'autres de manière immédiate, ce qui fait toujours de ce groupe le plus homogène des analyses de similarités.



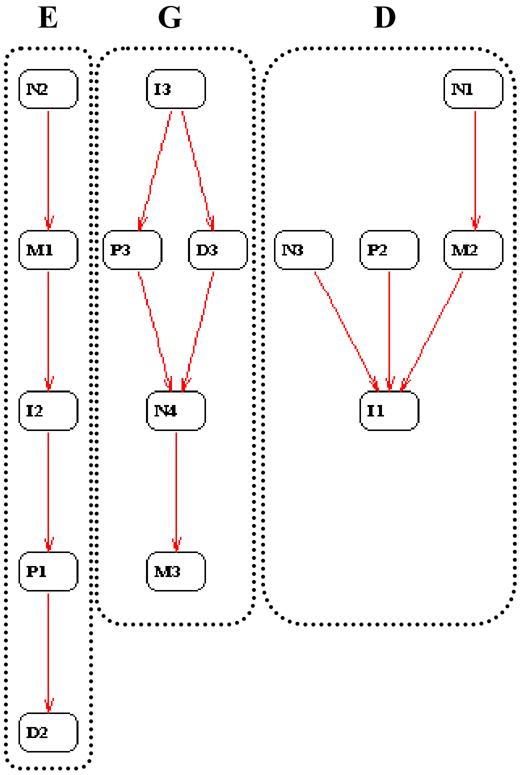
Cod	Var	Mod	Card
I3	Inter P	Non C	8
N4	Nat A	Non C	12
D3	Décis	Non C	10
P3	Calc P	Non C	10
M3	Manip	Non C	13

Tableau 46

A cause de l'absence de contexte, aucune interprétation de probabilité n'est envisageable (I3). Cette catégorie réunit généralement des exercices qui s'intéressent à faire exercer des aspects calculatoires du chapitre probabilité. Entre autres, le langage ensembliste, les axiomes de Kolmogorov et leurs propriétés. Cette classe est la moins intéressante à nos fins puisqu'elle exclue toute possibilité de lecture interprétative de la probabilité. Nous analyserons les mêmes exercices par un autre outil statistique, l'analyse implicative, qui nous permettra de trouver des règles quasi implicatives entre variables.

Graphe implicatif

Graphe Implicatif (seuil 0,85). Manuel Didier ES



Graphique 34

Le graphe implicatif correspond à un seuil de 0,85. Il présente trois groupes de quasi implications reliées, et à la différence des livres de texte de la filière Scientifique, ces trois graphiques sont sans intersection. Nous analyserons chacun de ces regroupements.

Groupe D

Ce groupe reprend les mêmes variables que le groupe fréquentiste A. L'implication N1 ⇒ M2 est de forte signification statistique, elle apparaît même à un seuil de 0,99, de même que M2 ⇒ I1. La première des deux rend compte du choix de ce manuel, comme tous d'ailleurs, de faire de la simulation une condition nécessaire au traitement d'objets de nature frequentiste, et cela parce qu'aucun manuel ne propose de contextes réels. La deuxième nous renseigne d'une autre implication : *s'il y a une simulation alors la probabilité calculée sera interprétée dans l'exercice.*

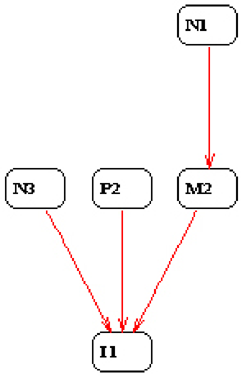


Tableau 47

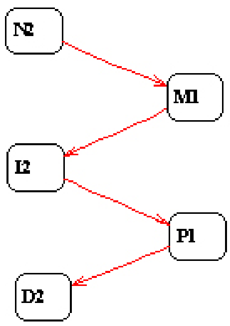
Cod	Var	Mod	Card
M2	Manip	Simul	11
I1	Inter P	Oui	12
N3	Nat A	Amb	9
N1	Nat A	Série	8
P2	Calc P	Impl	6

L'implication N3 ⇒ I1 qui apparaît pour la première fois à un seuil de 0,87, nous renseigne de situations dont les natures des objets alternent entre bayésienne et fréquentiste. Ce manuel de même qu'ailleurs, ne limite pas l'utilisation de la stabilisation à des situations de nature strictement fréquentistes. Il propose aussi des exercices dont le principe fréquentiste est une raison à croire pour évaluer numériquement la probabilité portée sur une épreuve générique.

Enfin, ce manuel reprend en grandes lignes, les principes dessinés par ceux de la filière Scientifique, un usage non explicite de l'intrication de signifiés (principe frequentiste), le traitement de la notion fréquentiste par des simulations, etc. Nous analyserons un deuxième groupe du graphe implicatif du manuel Didier ES.

Groupe E

La variable qui définit le groupe est, comme toujours, la nature de l'objet sur lequel porte la probabilité (épreuve générique). De la même manière que pour les autres groupes E des manuels analysés, le calcul de leurs probabilités n'est pas soumis à interprétation (I2).

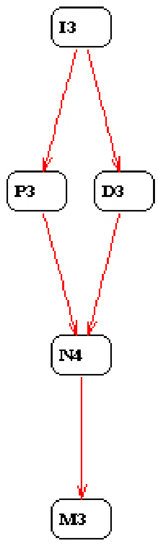


Cod	Var	Mod	Card
N2	Nat A	Génér	36
I2	Inter P	Non	45
M1	Manip	Evoq	41
P1	Calc P	Expl	49
D2	Décis	Non	55

Les quasi implications ne nous intéressent pas toutes sémantiquement. De ce groupe nous retenons la dernière de l'enchaînement (P1 ⇒D2), elle apparaît pour la première fois à un seuil de 0,98. C'est en fait sa contreposée que nous utiliserons lors de nos expérimentations : *si il y a une prise de décision alors le calcul de la probabilité n'est pas demandé explicitement.* Dans nos expérimentations, nous essayerons de placer la probabilité comme un outil de prise de décision, nous reviendrons plus tard sur ce sujet. Finalement, nous analyserons le troisième groupe de l'arbre implicatif du manuel Didier ES.

Groupe G

Cet ensemble de variables ne fait que confirmer l'existence d'un groupe d'exercices sans contexte où les questions posées ne concernent pas la probabilité en soi.



Cod	Var	Mod	Card
I3	Inter P	Non C	8
N4	Nat A	Non C	12
D3	Décis	Non C	10
P3	Calc P	Non C	10
M3	Manip	Non C	13

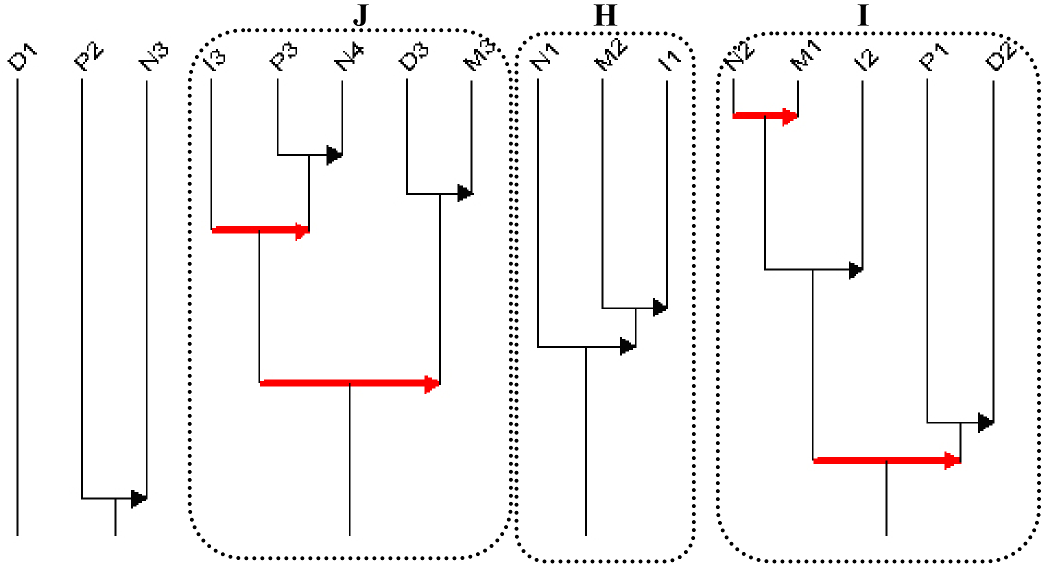
Tableau 48

Cet ensemble de règles apparaît dans le graphe à un seuil de 0,99, ce qui nous parle d'un profil consistant. De plus, si on se prête à l'exercice de descendre le seuil jusqu'à un niveau de 0,60 on observe que les groupes bayésiens et fréquentistes se mêlent progressivement au fur et à mesure que le seuil est de moins en moins restrictif, tandis que le groupe G reste toujours isolé. Ceci met en évidence la claire séparation de ces exercices au reste, ils ont l'exclusive fonction d'entraîner les élèves uniquement sur des aspects mathématiques de la probabilité. Nous continuons avec la troisième et dernière analyse du manuel Didier ES.

Arbre cohésitif

Des trois groupes retenus, le plus faible en nombre et qualité d'implications est le fréquentiste (H), suivi par le bayésien (I) et finalement le plus compact reste toujours le groupe d'exercices sans contexte (J) dont les raisons de son homogénéité ont été à plusieurs reprises expliquées. Nous commencerons par une description du groupe fréquentiste.

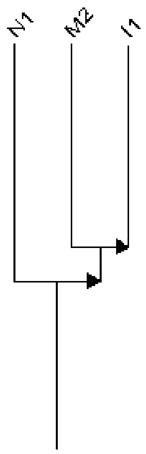
Arbre cohésitif. Manuel Didier ES



Graphique 36

Groupe H

Ce groupe réunit trois variables de sémantique importante, elles nous renseignent d'une part de l'utilisation de simulations conduisant à traiter la notion fréquentiste, et d'autre part de situations de nature ambiguë avec les caractéristiques déjà mentionnées.



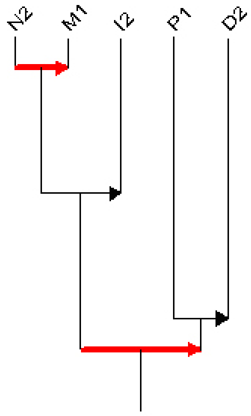
Cod	Var	Mod	Card
M2	Manip	Simul	11
I1	Inter P	Oui	12
N1	Nat A	Amb	8

Tableau 49

Le même schéma trouvé ailleurs se reproduit alors dans ce manuel, des exercices dont la notion fréquentiste n'est pas seulement la notion centrale mais aussi secondaire comme raison de croire pour évaluer une épreuve générique. Ce groupe se caractérise par l'utilisation de l'outil tableur (T3 t.a.r. 0,000158), pour poser entre 3 et 4 questions (Q2 t.a.r. 0,0749), par des contextes ludiques (C1 t.a.r. 0,116) et de cas élémentaires équiprobables (H1 t.a.r. 0,25). Enfin, des caractéristiques relativement similaires à celles des groupes fréquentistes des manuels de la filière Scientifique.

Groupe I

Ce groupe bayésien est plus nombreux que le fréquentiste tant en règles implicatives qu'en nombre d'exercices concernés, passant d'un groupe optimal de 9 exercices pour la classe fréquentiste à un de cardinal 36 pour cette classe bayésienne.



Cod	Var	Mod	Card
I2	Inter P	Non	45
M1	Manip	Evoq	41
N2	Nat A	Génér	36
D2	Décis	Non	55
P1	Calc P	Expl	49

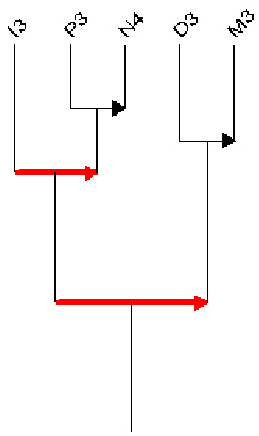
Tableau 50

Et tel que trouvé ailleurs, ce manuel propose des exercices de nature générique qui ne sont qu'évoqués ; ceux pour lesquels la probabilité calculée n'est pas soumise à une interprétation, en se focalisant plutôt sur des éléments du calcul de la probabilité que finalement n'interviendront pas pour agir devant l'incertain.

Cette classe se caractérise par le grand nombre de questions posées par exercice (Q3 t.a.r. 0,00108) et à la différence d'autres manuels, les énoncés sont de taille considérable (L2 t.a.r. 0,0755) dont la place de tâches ensemblistes est centrale (T1 t.a.r. 0,00172 et H1 t.a.r. 0,0188). Enfin, nous décrirons le dernier rassemblement de l'arbre cohésitif, le groupe qui correspond aux exercices sans contexte.

Groupe J

Cette classe, la plus homogène des trois, et comparable en nombre d'exercices à la fréquentiste, elle réunit tous les exercices manquant de contexte ou qui ne s'intéressent pas à la probabilité. Le manuel Didier ES reprend donc cette formule abandonnée par Bréal ES et avec une intensité assez considérable.



Cod	Var	Mod	Card
I3	Inter P	Non C	8
N4	Nat A	Non C	12
D3	Décis	Non C	10
P3	Calc P	Non C	10
M3	Manip	Non C	13

Tableau 51

Les règles et métarègles de cette classe sont les moins significatives à nos fins. En effet, le manque de contexte enlève toute possibilité d'analyser une interprétation associée à la probabilité. De ce groupe donc, nous n'analyserons pas la configuration de ses quasi-implications. Par la suite, nous résumerons donc l'étude sur le manuel Didier de la filière ES.

Groupe fréquentiste

Le groupe frequentiste a été présent chez le manuel Didier ES lors des trois analyses effectuées. Nous lui avons associés les mêmes noms que chez les autres manuels : Groupe A (Analyse de similarités), Groupes D (Analyse implicative) et Groupe H (Analyse cohésitive).

Ce groupe, de même qu'ailleurs, reste sensiblement moins nombreux que le bayésien, et en comparaison avec son homonyme de la même filière, il ne varie pas significativement en termes proportionnels.

A la différence du manuel Bréal ES, la maison Didier s'appuie plus fortement sur l'utilisation des outils informatiques, et c'est en fait par des simulations que les énoncés construisent les objets de nature fréquentiste. Il s'est crée ainsi une sorte de relation de dépendance de l'objet "série infinie" par rapport à l'ordinateur, en d'autres termes, les objets de nature fréquentiste ne sont traités qu'avec l'ordinateur, aucune autre possibilité pour la série ne s'étant présentée dans ces exercices.

De même que chez les autres manuels, cette classe inclut les objets de nature ambiguë, cette modalité nous renseigne, entre autres, de la présence d'objets de nature bayésienne (épreuve générique) qui sont évalués par le principe fréquentiste. Dans cette classe donc, se trouvent non seulement les exercices fréquentistes proprement dits mais aussi ceux qui, en évoquant la fréquence, demandent d'évaluer une probabilité qui porte sur un objet de nature bayésienne.

Les exercices de ce groupe, proposent en général des contextes ludiques dont le dénombrement en cas élémentaires équiprobables est facilement admissible. Ces exercices admettent donc une double voie d'évaluation de la probabilité, par le principe d'équipossibilité ou par la stabilisation de fréquences, l'une *a priori* l'autre a *posteriori*. Leur utilisation étant indiscriminée, l'exercice sollicite en général l'évaluation d'une probabilité sur une épreuve générique par le principe d'équipossibilité (qui ne peut être qu'*a priori*) et puis de la valider *a posteriori* par la stabilisation de fréquences.

Groupe bayésien

De même que le groupe fréquentiste, cette classe récupère la netteté perdue lors de l'analyse du manuel Bréal ES. En effet, chez Didier ES le groupe bayésien réunit les variables typiquement trouvées chez les manuels de la filière S. Et cela dans les trois analyses effectuées : parmi ces variables, les plus significatives concernent la nature de l'objet sur lequel porte la probabilité (épreuve générique), la non demande d'interprétation des calculs sollicités et la fixation sur des aspects calculatoires de la probabilité, en évitant toute demande de prise de décision à partir de la probabilité.

Ce groupe est toujours le plus nombreux et centre son attention sur les tâches calculatoires à partir de notions ensemblistes ou des axiomes de Kolmogorov, ses exercices permettant de faire exercer une ample gamme de techniques de calcul sans interpeller sur les signifiés des objets calculés.

Groupe sans contexte

Ce groupe aussi s'est vu manifesté lors des trois analyses effectuées avec le logiciel CHIC : Groupe C (Analyse de similarités), Groupe G (Analyse implicatif) et Groupe J (Analyse cohésitive).

Les trois groupes retenus (fréquentiste, bayésien et sans contexte) conforment une sorte d'échelle de graduation pour l'interprétation de la probabilité : le premier fait une référence explicite au signifié fréquentiste, même en situations ambiguës, le deuxième apporte un contexte bayésien mais ne s'intéresse pas au signifié de la probabilité, et le troisième rescinde le contexte, ou même l'objet probabilité, pour se consacrer entièrement et sans détours à des aspects calculatoires du chapitre Probabilités. Par la suite nous présenterons nos conclusions de l'analyse de ces quatre manuels des filières S et ES de la classe de Première du lycée français.

2.9 Conclusions

Dans notre premier chapitre nous nous sommes focalisé sur des aspects épistémologiques de la probabilité. A partir de la lecture d'acteurs clefs nous avons repéré tout au long de l'histoire de cette notion deux grands groupes interprétatifs, l'un bayésien, l'autre fréquentiste. En même temps cette lecture nous a permis de préciser un ensemble d'éléments caractéristiques de chacun de ces deux groupes interprétatifs. Avec ces éléments caractéristiques, nous avons cherché à nous outiller de critères objectifs nous permettant d'identifier le plus clairement possible la notion de la probabilité associée à un contexte donné.

Dans ce deuxième chapitre nous nous sommes servi précisément de ces éléments pour déceler la probabilité sous-jacente aux espaces potentiels de travail que représentent les exercices de manuels. Nous avons retenus quelques-uns de ces éléments caractéristiques sous la forme de variables et nous les avons intégrés à d'autres avec un double propos : d'une part pour confirmer l'incontournable dualité de la probabilité, et d'autre part pour nous informer des types de problèmes auxquels sont confrontés les élèves de lycée. Ce dernier aux fins de nos expérimentations en BTS que nous analyserons dans le troisième chapitre.

Le traitement effectué avec le logiciel CHIC des trente huit variables retenues dans les analyses des quatre manuels de la classe de Première nous a fournit un panorama assez clair du type d'exercices que les livres de textes ont tendance à proposer. Nous préciserons nos conclusions de l'analyse de ces quatre manuels de lycée français.

De la dualité de la probabilité

Dans notre premier chapitre et en nous basant sur notre enquête épistémologique nous avons postulé le caractère incontournable de la dualité de la probabilité. En d'autres termes, nous présupposions que toute tentative de découpage et d'élimination d'une des interprétations serait condamnée à un échec et que, et pour des raisons diverses, la dualité se manifesterait indéfectiblement dans les exercices, surmontant ainsi l'intention des rédacteurs des documents officiels et ceux des manuels de la réduire à l'approche fréquentiste.

La confirmation de notre hypothèse sur les exercices des manuels s'est centrée sur une variable : la nature de l'objet sur lequel porte la probabilité. Dans nos analyses des quatre manuels nous avons trouvé trois groupes d'exercices : bayesiens, fréquentistes et sans contexte. De cette manière l'hypothèse d'incontournable dualité est confirmée.

D'ailleurs les variables autres que la nature de l'objet ne sont pas moins importantes, elles nous renseignent sur des voies entreprises par les manuels pour se sortir de la contradiction consistant à considérer la probabilité comme fréquentiste, tandis que dans les faits les exercices sont, de nature bayésienne pour les uns, et fréquentiste pour les autres.

L'ensemble de variables retenues nous a permis d'identifier différents groupes d'exercices, ces groupes se sont manifesté plus ou moins clairement sur les trois analyses effectuées. Ainsi chaque analyse est un complément et une confirmation de l'autre. Nous résumerons les trois profils d'exercices trouvés dans ces analyses, un fréquentiste, un bayésien et un dernier sans contexte.

Groupe fréquentiste

Ce groupe d'exercices se constitue autour de la variable N1 (Nature d'A : série) et éventuellement de la variable N3 (Nature d'A : ambiguë). Le Tableau 52 est une reproduction du Tableau 17. Il nous montre une très faible présence d'exercices strictement fréquentistes, variant les pourcentages dès 1,4 % chez Nathan S à 12,3 % chez Didier ES.

Distribution d'exercices par nature d'objet à probabiliser (effectifs et pourcentage)

Manuel	Nature de A								Total
	Série		Epreuve générique		Ambiguë		Non correspon		
BELIN S	4	3,8%	36	34,6%	18	17,3%	46	44,2%	104
NATHAN S	1	1,4%	39	53,4%	22	30,1%	11	15,1%	73
BREAL ES	1	3,2%	22	71,0%	8	25,8%	0	0,0%	31
DIDIER ES	8	12,3%	36	55,4%	9	13,8%	12	18,5%	65

Tableau 52

Les exercices de nature strictement fréquentiste sont donc les moins nombreux ; ils semblent difficiles à construire au moins par rapport aux épreuves génériques. L'élaboration d'un contexte menant l'élève à se représenter mentalement l'objet série infinie est en effet plus coûteuse que celle d'une seule épreuve. L'objet série infinie nécessite un contexte où les reproductions sous les mêmes conditions soient facilement admissibles. Sans cette condition, la stabilisation de fréquences n'est pas garantie, et de plus si le contexte est relativement nouveau pour les élèves, il faut explorer la convergence de la proportion dans l'échantillon et cela implique des données.

Les données sont les grandes absentes des exercices analysés, nous parlons de données réelles, parce que tous les manuels ont opté pour reemplacer la réalité par des simulations. Ce choix n'est pas anodin, en France presque toutes les voies convergent pour une insertion des TICE, et les chapitres Statistique et Probabilité en semblent une bonne opportunité.

Bref, dans les manuels observés, la notion fréquentiste trouve un seul moyen pour se manifester, les simulations ; reste à savoir si ces dernières sont tant mobilisatrices que la réalité elle-même et si cette substitution est acceptée par les élèves pour transposer les conclusions vers la réalité de référence.

D'ailleurs, les contextes simulés du groupe fréquentiste traitent de cas élémentaires équiprobables, cette hypothèse étant le point de départ à l'élaboration des simulations. Sans elle les élèves ne pourraient pas les construire, tâche qui certes est fréquemment demandée dans ce genre de problèmes.

L'équipossibilité de cas élémentaires remplit des fonctions très importantes, elle permet par exemple de relier les deux interprétations de la probabilité. Ce principe peut être utilisé tant pour l'évaluation d'une épreuve générique que pour une évaluation *a priori* de la convergence de la série. Plusieurs exercices en profitent et bouclent dans un cycle les deux notions en demandant d'évaluer la probabilité d'une épreuve générique par la fréquence d'apparition, d'autres demandant l'évaluation de l'épreuve générique par le principe d'équipossibilité pour ensuite la valider par la fréquence. Cette dernière estimée par la proportion d'un échantillon provenant d'une simulation dont les cas élémentaires sont équiprobables, nous avons caractérisé ces exercices comme de nature ambiguë.

Le groupe fréquentiste donc n'est pas exclusif des exercices de nature purement fréquentiste, à cette catégorie arrivent ceux de la variable N3 (nature ambiguë) qui réunit ceux qui font évaluer la probabilité d'une épreuve générique par la proportion d'apparition du phénomène sur un grand nombre de reproductions, ces cas constituent une preuve de l'incontournable dualité de la probabilité et des voies entreprises par les manuels pour la traiter.

Néanmoins, toutes les épreuves génériques ne sont pas évaluées par le principe fréquentiste, bien au contraire, elles sont les moins nombreuses, dans son majorité, elles le sont par le principe d'équipossibilité, nous avons caractérisé ce groupe comme bayésien, par la suite nous résumerons les principales caractéristiques de cette classe.

Groupe Bayésien

Ce groupe représente une masse très importante d'exercices, ils fluctuent entre 34,6% (Belin S) et 71% (Bréal ES) d'exercices par manuel. A la différence de la classe précédente dont la fréquence d'apparition d'un phénomène devient pour quelques exercices source d'évaluation numérique d'une épreuve générique, ceux du groupe bayésien le font par le principe d'équipossibilité.

Pour ce groupe, les probabilités portent sur une caractéristique inconnue d'un phénomène unitaire ou événement simple, même si sa reproduction est toujours possible. Quelques cas sont plus évidents à cerner que d'autres, mais tous les exercices de ce groupe se ressemblent pour ces deux caractéristiques, l'objet étant de nature épreuve générique et évalué par le principe d'équipossibilité.

Néanmoins, la nature de l'objet et son signifié n'ont pas d'intérêt pour ces exercices, toute l'attention se centre sur les aspects calculatoires de la probabilité. De cette autre manière donc les manuels résolvent le conflit d'une dualité incontournable non reconnue, lorsque l'épreuve est de caractéristiques bayésiennes, et évaluable par le principe d'équipossibilité et qu'il n'y a pas de traces dans l'énoncé de la fréquence d'apparition du phénomène, les questions se cernent autour des calculs, techniques de dénombrement, etc. Le signifié des calculs effectués reste à la seule compréhension de l'élève, une sorte de construction personnelle qui ne devient pas objet à partager ou à institutionnaliser en classe.

Dans ce groupe, la variable P1 (demande explicite du calcul de probabilité) apparaît à plusieurs reprises dans nos analyses : d'une part elle met en évidence la direction vers laquelle s'orientent ces exercices et l'importance exclusive attribuée aux aspects calculatoires de la probabilité, et d'autre part elle nous renseigne d'une notion qui ne parvient pas à trouver une place en tant qu'outil de résolution d'un problème. En d'autres termes, la probabilité, une fois calculée, ne sert à rien ou en tout cas les exercices ne s'intéressent à ses possibles utilités. Nous soutenons que cette demande explicite récurrente est conséquence de la non prise en compte de son caractère d'outil.

Le rôle de cette notion comme outil visant une prise de décision rationnelle dans un contexte d'incertitude n'a été traité qu'exceptionnellement dans ces manuels. Les rédacteurs de ces exercices sollicitent le calcul de la probabilité explicitement ou par la demande de l'espérance mathématique. Pour ce dernier cas, c'est elle qui est demandée de manière explicite. Ni l'une ni l'autre de ces notions sont généralement interprétées, la structure de ces exercices pouvant être résumée comme suit : présentation de l'ensemble fini de référence, description d'un contexte donnant un cadre d'incertitude au résultat de l'action à entreprendre ou entreprise, demande explicite du calcul de la probabilité ou de l'espérance, fin de l'exercice.

Nous avons observé que les deux manuels de la filière S introduisent la notion d'espérance mathématique, et à la différence de son émergence à l'époque de Fermat et Pascal où les deux notions sont des outils pour fonder des décisions rationnelles, les choses se passent autrement dans les manuels : ni l'une ni l'autre n'auront cette fonction décisionnelle.

Pour le cas de l'espérance mathématique, les exercices se contenteront d'aspects déterministes de cette formule tel que par exemple le calcul d'une valeur de la variable ou d'une probabilité pour qu'un jeu soit équitable; jamais par exemple la maximisation de l'espérance mathématique devient un critère de décision.

Paradoxalement, les exercices du chapitre Probabilités ont la tendance à se circonscrire à des aspects déterministes de la probabilité : d'une part le langage utilisé donne l'impression de l'existence de résultats plausibles d'être évalués dans la dichotomie du vrai et faux (prouver, calculer, déterminer, déduire…) et d'autre part les évaluations des probabilités sont toujours plausibles qu'à partir d'un seul et unique résultat. En d'autres termes, lorsque une épreuve générique est évaluée numériquement par un ensemble fini de référence tel que celui de cas favorables parmi les possibles, jamais n'apparaît une option alternative de référence. Rappelons à cet effet les idées de de Finetti sur la probabilité subjective ou le cas de Camille que nous avons brièvement présenté dans le Chapitre I pour montrer comment différents ensembles de référence peuvent devenir source d'évaluation d'une probabilité. Ce "déterminisme" se soutient grâce à l'évocation sans exception d'un seul ensemble de référence.

Le concept d'épreuve générique en tant qu'événement simple ou de réalisation unitaire sera repris dans le dernier chapitre lors de nos conclusions générales et perspectives pour discuter sur le possible rôle des conceptions des chercheurs dans l'étude des conceptions des élèves. En effet, quelques chercheurs semblent avoir la même approche que ces manuels, ils semblent implicitement adhérer à l'idée qu'une épreuve générique n'admet qu'une seule évaluation numérique et décrivent comme erronées les réponses qui ne se correspondent pas au principe d'équipossibilité. Nous reviendrons sur ce sujet dans notre quatrième chapitre.

Bref, le groupe bayésien de ces manuels pourrait être décrit, sauf à quelques exceptions près, comme un ensemble d'exercices de nature épreuve générique, évaluable par le principe d'équipossibilité dont l'attention se fixe sur des techniques de calcul dans lesquelles le contexte incertain ne sert que comme cadre nécessaire à la notion de probabilité.

Il y a un troisième groupe, celui qui réunit les exercices sans contexte ou qui ne s'intéressent pas à la probabilité mais à des notions connexes telles que dénombrements, axiomes, etc. Nous décrirons très brièvement cette classe.

Groupe sans contexte

Ce groupe est le dernier en dégradation de l'interprétation de la probabilité, en effet, le groupe fréquentiste se trouve dans un extrême traitant une interprétation par un contexte associé, même si elle s'opère parfois de manière ambiguë ; puis nous trouvons le groupe bayésien dans lequel il y a une notion sous-jacente par le contexte évoqué mais sans reconnaissance institutionnelle, et finalement se trouve ce groupe pour lequel l'absence de contexte empêche toute signification possible.

Ces exercices varient en proportion allant de 0 % pour Bréal ES jusqu'à 44,2% chez Belin S. Par les caractéristiques nettement abstraites de ses exercices, cette classe empêche tout type d'interprétation, au moins celle qu'on pourrait se construire à partir du contexte du problème. Ainsi ils se focalisent sur des tâches techniques d'étapes précises du calcul d'une probabilité ou sur les axiomes de Kolmogorov qui ne requièrent d'aucune contextualisation pour les effectuer.

Cette classe ne présente pas un intérêt majeur à notre sujet de recherche autre que celui de souligner sa forte présence dans un des quatre manuels. De l'étude de ces quatre manuels nous retenons d'autres conclusions que nous détaillons par la suite.

La probabilité en tant qu'outil de prise de décision

Dans nos analyses nous avons introduit la variable D1 tendant à identifier les exercices dont la valeur d'une probabilité devient un argument de prise de décision. Notre recherche épistémologique nous a permis de repérer l'importance de son rôle comme élément de prise de décision. Des son émergence même elle est présentée comme un critère décisionnel, un exemple est constitué les lettres entre Pascal et Fermat dont pour décider sur des stratégies de jeu ou sur le convenable ou pas de la croyance en Dieu, ont fait recours à la probabilité.

Nos analyses nous renseignent qu'un seul manuel propose aux élèves de prendre des décisions. Environ 7% des exercices du manuel Bréal ES demandent aux élèves d'agir prenant plus ou moins en compte une probabilité calculée, le reste des manuels n'ayant pas franchi cette ligne et placent la probabilité comme un élément descriptif du cadre donnée par le contexte, jamais comme un outil pour agir face à l'incertain.

Nous pensons qu'il y a une relation de dépendance entre la demande explicite de la probabilité et la non prise de décision, cela a été reflété à plusieurs reprises dans les analyses par la quasi implication D2⇒P1 (si il n'y a pas de décision à entreprendre, alors le calcul de la probabilité est demandé explicitement). En d'autres termes, le caractère explicite de la demande de probabilité pourrait s'expliquer par la non demande de prise de décision. Dans nos expérimentations nous essayerons de profiter de cette dimension pour que la probabilité devienne un outil de résolution et qu'elle surgisse ainsi comme réponse à une nécessité.

Ainsi, la demande d'une prise de décision forcerait à habiller de signifié un calcul de probabilité. Rappelons en plus qu'interprétation de probabilité et prise de décision partagent le même type de registre de représentation, celui de la langue écrite et orale. Nous nous servirons de ces conclusions pour nos expérimentations en BTS.

Sur le genre d'utilité de la probabilité

Parmi les variables retenues nous avons considéré le type de contexte qui encadre un exercice, après plusieurs regroupements dus au faible nombre d'effectifs, nous avons proposé cinq modalités pour la variable type de contexte : Jeux, Trav-quot, Math, Autre et Non corresponde, le Tableau 53 est une reproduction du
Tableau 11.

Distribution d'exercices par type de contexte (effectifs et pourcentage)

	Type contexte										
Manuel	urnes-jeux-cartes		Clas-Trav-Quot		Math		Autre		Sans context		Total
BELIN 1S	*41*	*39,4%*	*10*	*9,6%*	*13*	*12,5%*	*23*	*22,1%*	*17*	*16,3%*	104
NATHAN 1S	*48*	*65,8%*	*12*	*16,4%*	*4*	*5,5%*	*4*	*5,5%*	*5*	*6,8%*	73
BREAL ES	*16*	*51,6%*	*13*	*41,9%*	*0*	*0,0%*	*2*	*6,5%*	*0*	*0,0%*	31
DIDIER ES	*36*	*61,0%*	*5*	*8,5%*	*2*	*3,4%*	*6*	*10,2%*	*10*	*16,9%*	59

Tableau 53

Dans ce tableau nous observons une forte tendance à proposer des contextes ludiques tels que jeux de cartes, urnes ou dés. La deuxième modalité, sauf le manuel Bréal est plutôt consacrée à des contextes de la réalité quotidienne dans la classe dont le calcul de probabilité n'a pas une transcendance significative. Pour ces exercices il est habituel de solliciter par exemple la probabilité selon laquelle le prochain élève qui rentre dans la salle soit un garçon sachant que la classe est composée de 30 % de garçons.

Nous nous servons des conclusions de ces deux chapitres pour proposer trois situations-problèmes à des élèves de BTS en France. Par la suite, nous détaillerons dans le chapitre III les analyses *a priori,* et *a posteriori* ainsi que les conclusions de ces expérimentations.

Chapitre III : Expérimentations en BTS

4.1 Introduction

Dans ce troisième chapitre nous présentons nos expérimentations en BTS. Le premier chapitre nous a permis une étude d'un point de vue épistémologique et il reste notre référence, le deuxième fut une approximation de l'enseignement de la probabilité par la voie des manuels. Dans cette troisième partie nous développons les analyses de trois expérimentations réalisées dans un BTS de la Région Bretagne pendant l'année 2006-2007. Ces trois séances ont été conçues pour sensibiliser les élèves sur les interprétations de la probabilité.

Les deux premiers chapitres peuvent être décrits comme une sorte de suivi de l'évolution d'un concept qui, d'un statut savant à deux visages fini par en perd un dans les documents officiels et qui le récupère partiellement dans les manuels (au moins dans ceux observés).

Dans sa transposition didactique, la dualité de la probabilité surmonte donc les barrières des programmes, mais à un coût non négligeable, elle perd le statut d'objet à enseigner d'un de ces composants (le bayésien). En effet, dans les manuels observés, les situations bayésiennes sont présentes, mais les effets de la transposition didactique ont fait que cet objet reste caché aux élèves par un détournement de l'attention vers les procédés de calcul d'une probabilité.

Pour ce troisième chapitre notre cherchons à ce que, dans une classe la plus ordinaire possible, un professeur et ses élèves puissent à la fin d'une séance, institutionnaliser une des notions de la probabilité. Par institutionnalisation nous entendons ici une identification ou une caractérisation au moins partielle de chacune des interprétations de la probabilité.

Dans ce projet, les situations-problèmes visent à créer les conditions favorisant l'émergence des respectives interprétations de la probabilité. En d'autres termes, nous cherchons à ce que les situations-problèmes facilitent l'utilisation chez les élèves de respectives interprétations de la probabilité.

De cette manière, les éléments caractéristiques décelés lors du premier chapitre constituent une référence à l'élaboration de nos situations-problèmes, ces éléments deviennent des conditions nécessaires de nos situations-problèmes. Par exemple, pour faciliter l'émergence en classe de la notion bayésienne, nous prendrons en compte soit l'élément caractéristique épreuve générique soit l'élément caractéristique hypothèse selon notre intérêt pour ce genre d'objets. Du même pour la notion fréquentiste, pour expérimenter une situation-problème accord à cette approche, nous prendrons en compte, entre autres, la série infinie comme objet sur lequel il faut probabiliser.

Les éléments caractéristiques s'avèrent donc des conditions nécessaires à prendre en compte lors de l'élaboration de nos situations. Néanmoins, ces conditions ne sont pas les seules ; pour qu'un tel projet puisse se concrétiser, il nous est indispensable qu'existe une coordination avec l'enseignant, coordination qu'implique une harmonisation des projets des deux parts. Cette coordination par l'harmonisation de projets nous a mené à développer des analyses *a priori* quelque part particuliers, dont nous avons détaillé à l'enseignant non seulement les possibles tâches des élèves, mais aussi un possible déroulement. Pour cette synthèse, nous avons préféré garder cette analyse premièrement, afin de mettre en évidence notre approche pour ces problèmes, et deuxièmement, car elle aidera à comprendre certains choix de l'enseignant lors de l'expérimentation. En effet, l'enseignant, nous le verrons lors de l'analyse *a posteriori*, a resté attentif à notre projet et quelques démarche pour lui effectuées semblent s'expliquer pour son intérêt à respecter notre projet d'origine.

Pour faciliter donc l'émergence et l'institutionnalisation des deux interprétations de la probabilité nous avons envisagé premièrement la création de situations appropriées à chaque notion et un moment de coordination avec l'enseignant pour d'éventuelles modifications ou adaptations à nos propositions de séances. Ces deux parties seront décrites ensemble au fur et à mesure que nous décrivons les situations-problèmes proposées.

Nous développerons ce qui, couramment, en didactique des mathématiques française est connu comme une analyse *a priori*, puis l'analyse *a posteriori* de chacune des trois expérimentations pour finalement proposer quelques conclusions.

Les caractéristiques particulières du sujet font que la présentation des analyses *a priori* de ces trois situations n'est pas organisée en trois parties, dont une pour chaque expérimentation. Il y aura en fait une première partie qui décrira quelques conditions générales à satisfaire pour n'importe quelle situation concernant les interprétations de la probabilité, puis une autre pour les conditions tendant à faciliter l'émergence d'une seule des interprétations pour finaliser sur une troisième partie pour décrire les variables locales de chaque situation-problème.

Pour les analyses *a posteriori*, nous réunirons ces trois niveaux dans une même analyse, une pour chaque expérimentation. Nous commençons par les caractéristiques générales, communes à n'importe quelle situation concernant les interprétations de la probabilité.

4.1 Caractéristiques générales

Avant de commencer avec la description des caractéristiques générales, nous souhaitons formuler quelques remarques ; elles concernent des questions didactiques.

La didactique de la statistique, en tant que sujet de recherche est très récente, en particulier en France où elle est peu développée. A cette rareté d'antécédents de recherches en didactique de la statistique nous devons ajouter les particularités de notre sujet épistémologique de référence : cela dans son ensemble, nous oblige à rester attentifs et prudents devant l'application de cadres théoriques importés de la didactique des mathématiques pouvant anéantir le relief propre à cette discipline.

La dualité de la probabilité est un sujet particulier, elle semble interroger et enrichir en même temps les différentes approches de la didactique des mathématiques. Nous pourrions citer par exemple la dimension décisionnelle pour l'approche de la dialectique outil-objet (Douady, 1986) ou la prise en compte de la sémantique en plus de l'opératoire en approches sémiotiques (Duval, 1993).

Il nous semble que l'adhésion à un cadre théorique particulier de didactique des mathématiques (TSD, TAD, etc.) serait prématuré, vu l'état du développement de la didactique de la statistique et de la recherche sur la dualité de la probabilité en particulier.

Alors, si risque il y a d'une acéphalie de théorie dans notre recherche, c'est parce que d'une certaine manière nous avons cherché à rester prudents et parce que nous acceptons, que notre recherche traite d'un sujet aussi important qu'inexploré en didactique de la statistique.

D'ailleurs, si dans cette première partie nous évoquons fréquemment par exemple le concept de variable didactique, cela n'implique pas que nous plaçons notre recherche dans le cadre de la Théorie des situations didactiques (Brousseau, 1998). Le lecteur doit prendre l'emploi de telles idées tout simplement comme la préhension d'un concept utile à notre recherche et sans aucune connotation d'adhésion totale à de telles théories. Ayant précisé notre position devant les différentes approches en didactiques des mathématiques, nous développons les caractéristiques générales de nos situations-problèmes.

Pour construire nos situations-problèmes, nous avons considéré trois niveaux, un premier de type général, commun aux deux interprétations, un deuxième afin de les différencier, et un troisième local qui gère les variables de chaque situation en particulier. Le premier de ces niveaux peut être décrit comme un cadre général, quelque elle soit la notion de la probabilité visée. En effet, les deux interprétations ont des caractéristiques communes et dans ce niveau nous plaçons les conditions générales tendant à les faire émerger. Dans le deuxième niveau nous nous intéressons aux variables didactiques dont leur maîtrise tend à faciliter l'émergence d'une seule des interprétations en bloquant l'autre.

L'approche que nous proposons pour traiter la dualité de la probabilité pourrait se comparer à celle d'une mise en scène théâtrale, dont deux acteurs (les interprétations de la probabilité) habillés avec un même costume jouent toujours ensemble (dualité), et un public (les élèves) qui doit les identifier. Ainsi, pour faciliter l'identification, le metteur en scène se sert d'un jeu de lumières sur le scénario afin de cacher provisoirement l'un des acteurs pour que l'autre devienne protagoniste.

Sous ce schéma, nos caractéristiques communes aux deux interprétations seraient représentées par la scène dans laquelle jouent les deux personnages et l'ensemble de variables tendant à les différentier serait le jeu de lumières. Nous développons les caractéristiques communes en commençant par le type de registre de représentation.

Type de registre de représentation

En effet, nos acteurs, les deux interprétations de la probabilité, ont l'habitude de se confondre pour jouer fréquemment ensemble et avec le même costume. En effet, la représentation symbolique P(A) est utilisée tant pour une attente à long terme que pour un degré de certitude. Un autre costume est aussi partagé, le registre numérique, tous les deux se servent de la même représentation symbolique. En d'autres termes ni P(A) ni son évaluation numérique tel que par exemple un huitième ne sont pas des moyens de représentation si riches comme pour discerner si un interlocuteur souhaite exprimer la fréquence avec laquelle il atteint quelque chose ou si c'est une mesure de sa certitude par rapport à une proposition donnée.

D'ailleurs, les représentations graphiques tel qu'un histogramme ou une courbe de Gauss se construisent à partir de ces registres, on ne doit pas s'attendre alors à ce qu'elles puissent récupérer le signifié de la notion de probabilité. En fait, toute représentation unifiée, commune aux deux interprétations, devient, en principe, incapable de fournir des éléments permettant d'exprimer le signifié d'une probabilité. Dans cet ensemble se trouvent les représentations symbolique, arithmétique et graphique.

Ces scènes (registres symboliques, graphiques, etc.) ne permettent pas aux interprétations de se manifester. Pour en trouver un sur lequel les interprétations puissent s'exprimer, nous avons besoin d'un registre plus riche, tel que le langage oral ou écrit. Ces moyens d'expression sont les seuls qui par leur diversité permettent d'éclairer le signifié d'une probabilité. Nos acteurs donc, devront jouer sur la scène du registre langagier.

Néanmoins, même ce registre reste le seul capable de disposer des éléments nécessaires à une telle précision d'idées, il présente aussi ses difficultés. Ce registre s'est fait écho de l'imbrication de ces personnages et a assigné le même mot (probabilité) aux deux interprétations. En effet, reflétant le caractère indissociable de la dualité, le langage la reproduit ; en français (probabilité), en espagnol (probabilidad) ou en anglais (probability), ces langues assignent le même signifiant linguistique à deux concepts bien différents rendant encore plus difficile leur distinction.

Toutefois, même si dans le registre langagier le signifiant est partagé, ce type de représentation reste le seul capable à mettre en évidence les différences entre les interprétations de la probabilité. En fait, le signifiant *probabilité* en commun n'est qu'une des difficultés pour différencier les interprétations dans le registre langagier. Une recherche de termes alternatifs conduit immédiatement au mot *chance* qui dans son usage courant semble reproduire le même inconvénient que le terme *probabilité* : il peut être utilisé pour se référer tant à la probabilité fréquentiste qu'à la bayésienne. Ce mot est donc un synonyme de la probabilité y compris dans sa dualité.

D'autres difficultés non négligeables sont à signaler, les unes proviennent des caractéristiques propres à ce registre, entre autres nous citons l'ambiguïté ou polysémie des mots, le manque de précision dans son usage et même l'attachement au contexte de leur sémantique (Barbier & Galatanu, 2000). Les autres concernent d'aspects didactiques tels que la monopolisation de la parole chez les enseignants, ce que rend difficile la reconnaissance des interprétations chez les élèves.

Néanmoins et en amont de toutes ces difficultés, il reste le seul sur lequel il est possible d'exprimer une interprétation de la probabilité, la seule solution étant le recours à des expressions synonymiques, nous reviendrons plus tard sur le sujet.

Ce registre fonctionne donc comme les planches de la scène sur laquelle nous chercherons à faire jouer nos deux acteurs. Sur ces planches se développeront aussi nos observations, c'est sur les échanges langagiers entre professeur et élèves et entre élèves eux-mêmes que nous centrerons nos analyses, et non pas autant sur les calculs, ces derniers ne nous renseignant pas sur la signification attribuée par un élève à une probabilité donnée.

Les prises de décision

Nous avons donc déterminé la base ou plateforme (planches de la scène) sur laquelle jouent les interprétations, le registre de la langue dans ses versions orale et écrite. D'ailleurs, dans le Chapitre II, lors des analyses des manuels, nous avons signalé l'importance de la dimension décisionnelle dans une approche enrichie de la dialectique outil-objet. D'une part, nous y avons observé que les interprétations et les prises de décisions partagent le même type de registre de représentation. Et d'autre part nous avons retenu la relation de dépendance entre une interprétation et une prise de décision, en d'autres termes, pour soutenir un argument décisionnel, il faut que le signifié de la probabilité soit précisé, sans celui-ci, la décision restant incomplète et inconsistante.

Nous allons nous servir de ces deux propriétés reliant les interprétations de la probabilité et les prises de décisions. La première de ces propriétés concerne donc le partage de registre, interprétations de la probabilité et prises de décision ne pouvant s'exprimer que dans le registre de la langue. La deuxième propriété est une relation de nécessité, pour soutenir une décision concernant la probabilité, il faut que cette dernière acquière une interprétation.

La combinaison de ces deux propriétés sera de grande utilité didactique lors de nos expérimentations. En demandant aux élèves de décider devant un contexte incertain nous introduisons indirectement le registre langagier. En interpellant les élèves sur les raisons de leurs décisions, nous forçons à expliciter le sens attribué à l'interprétation de la probabilité.

Dans notre schéma de représentation théâtral, la prise de décision devient ce que nous pourrions dénommer le rideau de fond de scène, ainsi non seulement nous traitons la question du type de registre approprié mais aussi celle de la probabilité comme un outil rationnel pour décider devant l'incertain. Les critères de décision, bien évidement, resterons simples, tels que : décider par le principe de maximisation de l'espérance ou celui de la de probabilité.

Autres registres de représentation

La dimension sémantique est fondamentale à la probabilité, mais elle n'est pas la seule. La probabilité se construit aussi sur une dimension numérique. Sans cet autre axe aucune méthode inférentielle telles que nous les connaissons aujourd'hui, ne serait possible. Cette dimension se développe sur des registres qui s'avèrent plus appropriés que la langue. Ainsi, nous trouvons la représentation numérique qui permet de préciser la valeur de la probabilité et la représentation symbolique sur laquelle s'expriment les règles de transformation et les propriétés de cohérence à la représentation numérique.

Ces deux registres, numérique et symbolique, sont autant indispensables que indissociables. Par exemple, sans le dernier, il n'y aurait pas une architecture pour les calculs de probabilité. Les axiomes de Kolmogorov fonctionnent comme des règles de cohérence aux représentations numériques. Ces axiomes, exprimés dans le registre symbolique, nous renseignent des normes de base à satisfaire tant pour une représentation numérique que pour ses transformations.

La représentation sur le registre numérique sera dans nos situations-problèmes, la conséquence de la nécessité de précision que le registre de la langue ne peut pas offrir. Pour provoquer le passage du registre dès la langue naturelle vers le numérique, il faudra que le premier devienne insuffisant. Pour cela les possibles solutions d'un problème seront si proches qu'elles demandent au registre symbolique une précision plus grande.

Les registres de représentation numérique et symbolique entreraient dans notre schéma de scénario théâtral en tant que nouvelles scènes pour ces personnages. La première a été le registre de la langue, sur cette scène jouent les interprétations, sur celle du registre numérique joue l'évaluation de la probabilité et sur la dernière, le registre symbolique, se déploient les critères de transformation.

La communication entre le registre langagier et le numérique est assurée par la nécessité de précision qui caractérise nos problèmes, le passage entre le registre numérique et le symbolique par la nécessité de cohérence, toutes les deux établissant une sorte de ponts entre planches (

Figure 14).

Registres de représentation de la probabilité

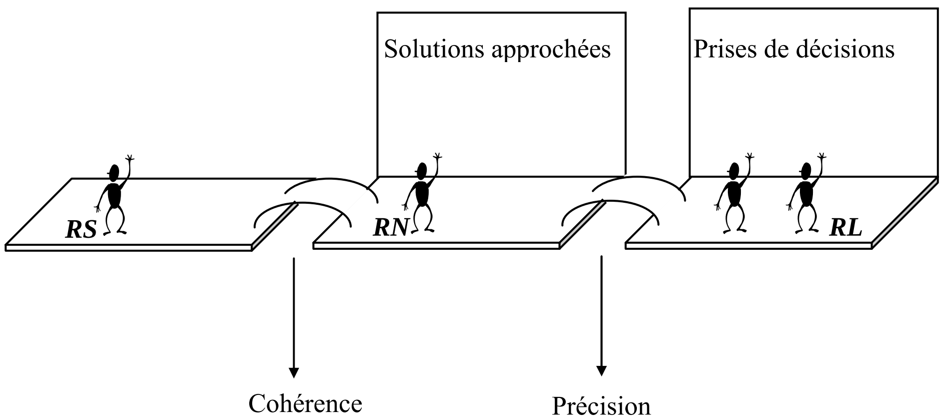


Figure 14

Contextes réels

Dans notre comparaison nous avons placé des scènes représentant des registres sémiotiques : la première a été pour le langagier, sur cette plate-forme jouent les interprétations de la probabilité, et un rideau de fond a été ajouté, symbolisant un contexte de prise de décision. Suivant cette métaphore, nous incorporons les contextes réels.

En effet, pour nos expérimentations et à la différence des manuels et des documents officiels, nous privilégions ce genre de contextes. Ils disposent de propriétés très utiles à la construction des interprétations de la probabilité. Nous précisons ces propriétés dont quelques-unes ont été déjà commentées lors des analyses des manuels.

Préservation de la qualité de l'information

La qualité de l'information pour les prises de décisions sollicitées trouve son point optimal dans les contextes réels, tout autre genre de contexte fournissant une qualité inférieure. Cette qualité contribue entre autres à :

- Ecarter les doutes que peuvent avoir les élèves sur certains choix de l'enseignant. Des choix qui pourraient être perçus comme une manipulation de la part de l'enseignant à des fins didactiques. Ces doutes pourraient discréditer les arguments soutenant les prises de décision.
- Assurer la valeur logique de certains événements dont dépendent les calculs de probabilités conditionnelles. Par exemple si les élèves doivent calculer $P_B(A)$, alors faire en sorte que la certitude sur la valeur logique de B ne soit pas mise en doute, cet item concernant la probabilité bayésienne.
- Créer les conditions d'admissibilité des hypothèses pour le modèle fréquentiste. Rappelons que, sauf une estimation par la convergence de l'échantillon, la probabilité fréquentiste est calculée sous l'admission d'un ensemble d'hypothèses. Il est donc souhaitable une certaine connaissance du système pour admettre les hypothèses, cette connaissance étant optimisée lors de la présence des objets.

Mobilisation de conceptions déterministes

Cette propriété a été déjà commentée lors des analyses des manuels. A l'occasion, nous avons exprimé nos réserves sur la capacité des contextes évoqués ou simulés à mobiliser des débats concernant de possibles relations de causalité dans les événements étudiés.

L'efficacité des contextes évoqués ou simulés pour remettre en question les conceptions déterministes ne nous semblent pas suffisamment étudiées. Le sujet mérite une recherche approfondie, néanmoins les simulations ou évocations ne nous semblent pas tant mobilisatrices qu'un contexte réel, ce dernier restant la source naturelle de questionnement aux modèles déterministes du monde.

Transparence des simulations

Cette propriété a aussi été commentée dans le deuxième chapitre, elle concerne le problème de l'acceptation d'une simulation comme équivalente à la situation réelle de référence. Nous avons confirmé le doute sur une hypothétique transparence lors d'une expérimentation exploratoire et antérieure à celles qui nous occupent dans cette synthèse. A l'occasion nous avons observé deux phénomènes pouvant être dénommés comme de « non transparence » d'une simulation pour l'un, et comme de « persistance déterministe » pour l'autre. Ces phénomènes, nous les avons constatés lors d'une activité proposée à une enseignante de terminale S à Paris en printemps de l'année 2005. La situation en question, de caractère ludique, se déroulait d'abord sur un contexte réel et puis sur un support informatique.

Dans la première séance les élèves, travaillant en couples, avaient produit leurs propres données, en lançant deux dés. Trois jours plus tard, dans une deuxième séance les données se étaient générées sur ordinateur.

Avec ce jeu, nous avions eu l'intention d'introduire quelques principes du test de X^2 (probabilité fréquentiste). Les élèves devaient distribuer 36 pièces de monnaie parmi tous les résultats possibles de la somme des faces de deux dés (2, 3, …, 11, 12). Puis, ils devaient lancer 20 fois leur paire de dés en récupérant à chaque lancer une pièce de monnaie de la boîte correspondante au résultat obtenu de la somme des faces visibles des deux dés. Pour gagner, il fallait récupérer plus de pièces que l'opposant. Nous avons demandé aux élèves de proposer une stratégie gagnante à ce jeu.

De cette expérimentation nous souhaitons retenir ici deux phénomènes observés. Le premier a été la manifestation persistante de conceptions déterministes chez quelques élèves. Pour ces élèves un ordre supérieur régissait les lancers. Ces réactions ont eu lieu lors de la première séance, quand la situation se déroulait dans un contexte réel. Le deuxième concerne le refus de la presque totalité des élèves à accepter la simulation de la deuxième séance comme une continuation de la manipulation réelle de la première. Cette réaction presque unanime nous a été explicitement signalée par quelques élèves et soutenue par le reste de la classe.

Cette expérimentation nous a confirmé nos doutes sur l'efficacité à faire émerger les conceptions déterministes et sur la transparence des simulations. En résumé, il nous semble que les simulations doivent être utilisées avec précaution, principalement lorsque la notion de probabilité est en train de se construire. Les simulations, bien qu'elles aient leurs avantages sur certains points, peuvent aussi par un usage abusif, introduire des difficultés dans l'apprentissage de notions de statistique : parmi elles, celle de devenir des réalités en soi dont les conclusions pourraient ne pas être transposables à celle de référence, empêchant toute autre application que la virtuelle. Dans nos situations-problèmes, nous avons donc pris une position prudente devant ce genre de contextes, en choisissant la manipulation réelle à la place d'une simulation.

Fixation de la nature de l'objet sur lequel porte la probabilité

Cette autre propriété a été elle aussi commentée lors des analyses des manuels. Elle concerne principalement les épreuves génériques. Rappelons quelques caractéristiques de ce genre d'objets ; ce sont des événements unitaires (par leur cardinalité), perçus comme un parmi d'autres, et susceptibles de reproduction sous les mêmes conditions. Pour l'évaluation de sa probabilité, deux principes sont généralement utilisés, l'un connu comme d'équipossibilité, l'autre comme fréquentiste.

Ces épreuves sont facilement convertibles en fréquentistes. Il suffit tout simplement de modifier l'objet sur lequel porte la probabilité, passant ainsi d'une épreuve générique à la série infinie. L'évaluation numérique ne changeant pas, la seule altération reste la nature de l'objet. Altération qui reste inaperçue si l'attention se concentre sur les calculs.

Cette potentielle conversion d'objet dans les épreuves génériques est facilitée d'une part par l'intérêt exclusif mis sur les calculs et d'autre part par le caractère évocateur du contexte. En effet, les calculs, invariables à ce niveau, n'ont pas besoin de la précision de la nature de l'objet. L'autre condition, le caractère évocateur, a été commentée lors des analyses des manuels, il suffit d'un simple ajustement rhétorique pour que l'objet passe du domaine bayésien vers le fréquentiste.

Afin de mieux cerner la nature de ces objets et de la valeur logique en question, nous proposons la présence de l'artefact qui produit l'épreuve. En d'autres termes, que le contexte se matérialise en classe afin de permettre, le cas échéant, de confirmer sa nature par évidences factuelles.

La Figure 15 est la représentation complète des caractéristiques communes de nos situations-problèmes. Elle pourrait être étendue en ajoutant d'autres registres tel que le graphique, néanmoins, l'architecture du dessin resterait invariante. Le registre dans lequel les interprétations s'expriment est le langagier, il est le seul à permettre la différenciation des deux interprétations. Nos problèmes ont en commun l'utilisation de la probabilité pour une prise de décision, ainsi il devient rideau de fond. Une autre caractéristique partagée pour toutes les situations est la matérialisation des contextes. Les autres dimensions de la probabilité sont assurées par les registres numérique et symbolique, sur lesquels se développent respectivement l'évaluation numérique et les règles de transformation.

Représentation de caractéristiques communes aux deux interprétations de la probabilité

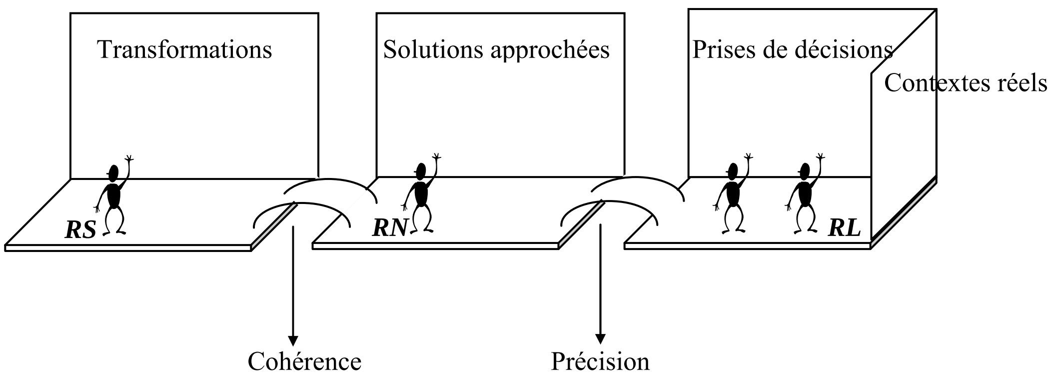


Figure 15

Ces caractéristiques, en tant que communes aux deux notions, n'ont pas vocation à trouver les différences entre une notion et l'autre, elles constituent un cadre général tendant à favoriser l'émergence de la probabilité en classe. Pour trouver des éléments différenciateurs, nous nous appuierons sur les éléments caractéristiques décrits dans le Chapitre I de cette synthèse.

4.1 Caractéristiques spécifiques

Les caractéristiques spécifiques nous permettent d'orienter la situation vers une approche fréquentiste ou bayésienne selon notre intérêt. Elles sont représentées dans notre mise en scène par un dispositif gérant les lumières du scénario, cherchant à cacher l'une pour rendre protagoniste à l'autre.

Ces caractéristiques spécifiques dérivent directement des éléments caractéristiques repérés dans le premier chapitre. Ils sont :

- La nature de l'objet sur lequel porte la probabilité
- Les valeurs logiques
- Les principes d'évaluation
- Le type de raisonnement

Nous avons récapitulé dans le Chapitre I (Conclusions) cet ensemble de différences, ci-contre, nous reproduisons le tableau les synthétisant.

Quatre différences entre les interprétations de la probabilité

		Interprétations			
		Fréquentiste		Bayésienne	
Différences	Objet sur le quel probabiliser	Série infinie		Epreuve générique	Hypothèse
	Valeurs logiques portés sur ces objets	Vrai- faux		Intervalle continue [0,1]	
	Type de raisonnement	Déduction		Référenciation	Abduction
	Quelques principes de quantification	Hypothèses du modèle	Equipossibilité (Ensemble fini)	Fréquentiste (ensemble infini)	Raison insuffisante

L'importance de ces éléments devient plus claire lorsqu'ils sont projetés sur les situations-problèmes. C'est dans sa matérialisation dans les situations qu'ils sont mieux appréciés. Nous présentons donc les analyses des situations expérimentées en classe de BTS. Les éléments caractéristiques seront développés sur chacune des situations.

4.1 Analyse de la première expérimentation

Nous avons expérimentés trois situations-problèmes, elles seront présentées dans leur ordre de réalisation. Les deux premières ont eu lieu en Novembre 2006, la dernière en Mars 2007. La classe était composée de 22 élèves dont une fille, tous âgés entre 18 et 21 ans. Au moment de l'expérimentation les élèves se trouvaient en première année d'un BTS électrotechnique. Nous décrivons la première expérimentation.

Le jeu des pièces de monnaie

Analyse *a priori*

Matériel :

Quatre pièces de monnaies par élève. Papier. Crayon.

Le jeu :

Chaque élève doit définir le nombre de pièces (k) avec lequel il jouera toute la partie. L'élève lance ses k pièces. Deux possibilités se présentent :

- Si toutes les k faces supérieures sont pile, l'élève accumule 10 points pour chaque pile.
- Si parmi les k faces supérieures il y a au moins une face, l'élève n'accumule pas de points.

L'élève qui arrive à accumuler 50 points dans le moindre nombre de coups gagne la partie. Question : Avec combien de pièces convient-t-il de jouer ? Justifier

Objectif pour l'expérimentation

En général, nos expérimentations visent à étudier les conditions et les possibilités de sensibilisation de la dualité de la probabilité dans une classe ordinaire de mathématiques. Pour cette expérimentation nous nous intéressons à celles de la notion fréquentiste.

Plus précisément, nous cherchons à :

- Confirmer si la notion de probabilité évoquée par professeurs et élèves correspond à celle prévue par la gestion d'éléments caractéristiques.
- Analyser les difficultés et obstacles rencontrés dans cette expérimentation en relation à notre intention de sensibilisation de la dualité de la probabilité dans une classe ordinaire.

L'analyse se réalisera sur les échanges langagiers produits lors de la séance et les échanges que nous avons eus avec l'enseignant avant la séance. Pour cela et en souhaitant perturber le moins possible la classe, nous avons placé une vidéo caméra au fond de la salle et distribué quatre magnétophones sur quelques tables. Les échanges avec l'enseignant, au nombre de six (environ deux heures en tout) se sont effectués sur un logiciel de communication par connexion Internet.

Pour montrer les enjeux du problème nous présentons un exemple illustratif. Admettons que l'élève A choisi de jouer avec 3 pièces de monnaie et l'élève B avec 1 pièce. L'élève A commence la partie et en lançant ses pièces il obtient pile, pile, face. Il n'accumule pas de points pour ce lancer. L'élève B en lançant sa pièce obtient pile, il accumule donc dix points. La Tableau 54 résume un possible déroulement.

Elève A	Lancer	*P P F*	*P P P*	*F F F*	*P F P*	*F P P*	*P P F*	*P P P*
	Points	0	30	0	0	0	0	30
	Total	**0**	**30**	**30**	**30**	**30**	**30**	**60**
Elève B	Lancer	*P*	*F*	*P*	*P*	*F*	*P*	*F*
	Points	10	0	10	10	0	10	0
	Total	**10**	**10**	**20**	**30**	**30**	**40**	**40**

Tableau 54

Après sept lancers, l'élève A est le premier à arriver à accumuler cinquante points. Il gagne la partie.

Solution

Critère de maximisation de l'espérance.

Soit X les points gagnés par lancer. En jouant avec :

Une pièce de monnaie, E(X) = $10\times\frac{1}{2}+0\times\frac{1}{2}=5$

Deux pièces de monnaie, E(X) = $20\times\frac{1}{4}+0\times\frac{3}{4}=5$

Trois pièces de monnaie, E(X) = $30\times\frac{1}{8}+0\times\frac{7}{8}=3{,}75$

Quatre pièces de monnaie, E(X) = $40\times\frac{1}{16}+0\times\frac{15}{16}=2{,}5$

Stratégie gagnante : jouer avec une ou deux pièces.

Déroulement :

Cette situation a été conçue comme la première du chapitre Probabilité, en d'autres termes, non seulement elle est notre première expérimentation, la séance en question sera la première du chapitre probabilité pour ces élèves. Nous commençons par quelques commentaires pour ensuite continuer avec l'analyse *a priori* du déroulement de cette première situation.

Nous avons admis une certaine maîtrise des calculs élémentaires de probabilité chez ces élèves, fruit de leur passage par le lycée. D'ailleurs, les analyses des manuels nous ont montré une forte présence de contextes ludiques dans les exercices. Pour cette première expérimentation nous avons choisi ce genre de contexte pour faciliter la récupération de notions anciennes (Robert & Rogalski, 2002). Néanmoins, bien que le contexte ludique leur soit habituel, le type de situation peut déstabiliser quelques élèves, dû à plusieurs nouveautés :

- Matérialisation du contexte
- Demande de prise de décision
- Problème à deux solutions optimales
- Utilisation du critère de maximisation de l'espérance

Probablement P doit réaliser plusieurs tirages afin de faire comprendre les règles du jeu aux élèves. Dans cette étape, l'enseignant éviterait les notations symboliques spécifiques et les références au mot probabilité afin de ne pas l'introduire de manière anticipée.

Bien que rare, quelques élèves peuvent être plus ou moins convaincus de pouvoir contrôler les résultats des tirages. Si un de ces élèves se trouve dans ce stage, il semble difficile un changement d'avis lors de cette première séance. Il faudra revenir sur cette persistance déterministe à plusieurs reprises. Par contre nous attendons pour la plupart de ces élèves leur reconnaissance de l'impossibilité de prédiction, principalement sur ce genre de contexte. Les situations de jeux simples (pièces de monnaie, dés, jetons, etc.) sont courantes dans la vie personnelle des élèves, nous admettons que pour ce genre de contextes, une approche indéterministe est relativement facile à admettre.

Il y a dans cette première situation l'intention de faire jouer quelques notions anciennes et d'introduire des nouvelles. L'ancien, est présent par un contexte de jeux, les pièces de monnaie et les calculs des probabilités dans des contextes ludiques. Le nouveau serait l'association du calcul de probabilité à la stabilisation de fréquences et l'utilisation de cette stabilisation comme notion fondamentale de l'argument décisionnel.

Nous déclinons notre analyse en cinq étapes, chacune étant caractérisée par des démarches spécifiques. Nous décrivons ces cinq phases en précisant les principales tâches des élèves et du professeur (P).

Premier étape : exploration du problème

Comme d'habitude la dualité de la probabilité est présente dans cet exercice, nous décrivons cette double présence.

Les deux interprétations de la probabilité

Bien que cette situation vise l'association entre la stabilisation de fréquences et la probabilité, la présence de la notion bayésienne est incontournable. Une des raisons est que personne ne conçoit naturellement jouer jusqu'à l'infini. Au contraire, un joueur s'intéresserait plutôt à choisir une stratégie lui permettant de gagner *la* partie qu'il va jouer. C'est un des problèmes majeurs de la notion fréquentiste, le contexte de son application est celui d'une réalisation répétée un grand nombre de fois sous les mêmes conditions.

L'intérêt de gagner une partie correspond à la probabilité bayésienne. En effet, en s'intéressant à la partie qui va commencer, le joueur centre son attention en ce que nous avons dénommé un cas générique. Mais à la différence des évaluations des cas génériques observés dans les manuels, celle de la probabilité de gagner une partie de ce jeu est complexe, la difficulté provenant de l'absence d'un ensemble de référence. Ainsi, par le manque de cet ensemble, la valeur numérique de cette probabilité bayésienne devient inaccessible à l'élève, et elle reste qualitative (Keynes, 1921), sans sa composante numérique.

Nous nous servons de son incomplétude pour bloquer le regard bayésien et centrer toute l'attention sur la notion fréquentiste. De cette manière la probabilité bayésienne reste cachée. Le rôle protagoniste est alors assumé par la notion fréquentiste, seule notion qui dans ce problème peut être évaluée numériquement, arrivant ainsi à remplir ses deux dimensions, la sémantique et l'opératoire.

La probabilité fréquentiste sera donc la notion centrale de ce problème, elle intervient dans le calcul de l'espérance mathématique, qui par le critère de sa maximisation devient argument du choix du nombre de pièces.

L'argumentation du choix du nombre de pièces peut ne pas être évidente pour les élèves, ils se verront tentés de choisir quatre pièces, cela leur rapporterait quarante points d'un seul coup, mais la rareté du phénomène ($P(pppp)=1/16$) rend ce choix peut efficace. Il y a donc dans cette situation une première phase de découverte de la nécessité de considérer de manière conjointe les points obtenus par lancer et sa rareté.

Nous faisons l'hypothèse que les élèves n'auront pas de grandes difficultés à admettre une stabilisation de fréquences pour chaque issue possible de ce jeu. Néanmoins, ils pourraient se voir confrontées à plusieurs nouveautés, parmi elles :

- Déterminer *a priori* la fréquence d'apparition.
- Intégrer de manière conjointe la rareté de l'issue aux points gagnés par lancer.
- Utiliser le critère de maximisation de l'espérance comme l'argument décisionnel.

Même si la stabilisation de fréquences nous semble relativement facile à admettre dans ce genre de contextes, nous prévoyons un temps consacré à son étude, qui serait, à en juger par les exercices des manuels, la première fois que ces élèves expérimentent de telles démarches.

La première étape de cette situation est donc d'exploration, d'appropriation du jeu, de découverte de la nécessité de combiner les points gagnés avec la fréquence, tout en travaillant avec des données réelles.

Deuxième étape : Rassemblement des données

La production des données contribue à installer l'objet sur lequel porte la probabilité. L'idée est de construire progressivement l'objet série infinie accompagné des valeurs logiques associées.

Cette première production de données doit rester restreinte en nombre, tout en permettant aux élèves de découvrir la nécessité de combiner points gagnés par lancer avec fréquence d'apparition. Cette limitation du nombre de lancers se doit principalement à notre objectif de souhaiter associer le calcul de probabilités (par des hypothèses admises) avec la stabilisation de fréquences. Si ces premières données étaient suffisantes aux élèves, ils ne ressentiraient pas la nécessité du calcul *a priori* de la probabilité.

Toutefois, l'échantillon produit par les élèves a des fonctions importantes, avec cet échantillon nous cherchons à emmener la discussion sur le *collective* (von Mises, 1928), visant la régularité de la série infinie au lieu de celle d'expériences personnelles. Du point de vue de l'approche fréquentiste, la régularité est une caractéristique de la série et indépendante des joueurs, et cela peut être induit en réunissant les données de tous les élèves ; nous cherchons ainsi à "neutraliser" les arguments déterministes (des élèves plus chanceux que d'autres, etc.).

A la différence de la première qui était exploratoire où d'approximation au problème, dans cette deuxième étape chaque élève doit laisser des traces de ses lancers afin de les traiter ensemble. Ces données seront réunies dans un tableau qui cherche à centrer l'attention sur les régularités du système. Combien de lancers faut-il faire ? Un seul par élève et par issue possible. Avec vingt élèves, le tableau semblerait à celui-ci :

	Tirages																			
Issues	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*
Un	10	10	10	0	0	10	0	0	0	10	10	0	10	0	0	10	10	0	0	10
Deux	20	0	0	0	0	0	20	20	20	0	0	0	0	0	20	0	20	0	20	0
Trois	0	0	0	0	30	0	0	0	0	30	0	0	0	0	30	0	0	0	30	0
Quatre	0	0	0	0	0	0	0	40	0	0	0	0	0	0	40	0	0	0	0	0

La réunion des données permettrait d'une part de détourner l'attention des données personnelles vers les tendances du jeu (l'objet série infinie), et d'autre part de reconnaître les deux variables du problème, points gagnés et fréquence d'apparition. Nous cherchons en fait à ce que cet ensemble reste restreint, pour motiver, entre autres, la recherche du calcul des probabilités, sous l'excuse didactique de trouver des moyens plus économiques que les lancers. La contrainte des données insuffisantes en nombre a donc l'intention d'amener les élèves à trouver le modèle de probabilité.

Avec les données au tableau, il y a un risque, que les élèves fixent leur attention sur elles, et qu'ils les utilisent comme argument de leur choix. Les élèves pourraient proposer par exemple, de faire la somme de points gagnés par ligne et déterminer une stratégie gagnante à partir du résultat de cette somme, ou de compter combien de zéros il y a dans chaque colonne, etc. Ainsi, les élèves pourraient prendre les informations du tableau comme définitives. Ce genre d'argumentation doit pouvoir être rapidement réfuté (à la limite pour P) : *« d'autres tirages donneraient un résultat différent… ».* P pourrait demander une nouvelle série de tirages pour en suite écrire ces nouveaux résultats sur le tableau, mais en effaçant la liste précédente.

Du débat supposé installé, il faut donc que la classe s'intéresse aux deux variables en jeu :

- la rareté de chaque issue possible
- le montant du gain

Troisième étape : Validation du critère de maximisation

Bien que les élèves puissent reconnaître l'interrelation entre les variables intervenantes, il nous semble difficile qu'ils s'engagent à calculer l'espérance de leur propre initiative, même s'ils disposent des connaissances nécessaires. L'inconvénient majeur, il nous semble, proviendrait du changement de paradigme, du déterministe vers l'indéterministe.

En effet, franchir le pas vers l'indéterminisme ne semble pas du tout évident. Les actions habituelles entreprises en classe de mathématiques s'appuient en partie sur la valeur logique (vrai-faux) des propositions. Tandis que celles basées sur des raisonnements inductifs ou abductifs le font entièrement sur l'enchaînement argumentatif, sans prise en compte de la valeur logique de la proposition en question, tout simplement parce qu'elle reste inconnue.

Par exemple, la valeur logique de la proposition *« je gagnerai la partie »* reste inconnue à l'élève, il devrait effectuer son choix par un enchaînement d'arguments connu comme critère de maximisation de l'espérance. Cet enchaînement d'arguments ne lui permet pas de prédire s'il gagnera ou non. Plus encore, il devrait continuer à choisir la même stratégie même s'il vient d'en perdre une. Stratégie qui ne se voit pas remise en cause par la valeur logique « faux » de la partie perdue. La stratégie ainsi garde sa validité par sa cohérence argumentative et non pas par une de ses réalisations.

Ceci est important parce que les élèves, au moment où nous leur demanderons de choisir une stratégie gagnante, disposeront de données qui pourraient être utilisés pour falsifier la stratégie gagnante. Habitués en classe de mathématiques à ce genre de démarches, les élèves pourraient se voir inhibés à proposer d'autres arguments.

La résolution de ce problème requiert un repositionnement de paradigme et un contrat didactique particulier, un où les contre-exemples et les certitudes disparaissent et laissent la place à des procédés argumentatifs acceptés par leur rationalité et non pas par leur falsifiabilité.

A la difficulté didactique qu'impliquent ces nouveaux critères de validation nous devons ajouter la récurrente présence de conventions dans les critères décisionnels en statistique. Des critères étant parfois perçus comme arbitraires ou subjectifs. Le critère de maximisation de l'espérance en est un exemple, un autre plus ou moins naturel à admettre correspond à la détermination de la région de rejet d'un test d'hypothèse fréquentiste établi à 5 % de probabilité.

Tous ces changements peuvent inhiber l'action des élèves. Il nous semble donc nécessaire de prévoir un blocage et prévenir l'enseignant de la possible nécessité de son intervention pour valider et/ou encourager des initiatives tendant à raisonner autrement. Dans ce sens, les interventions de l'enseignant du type : *« Si on joue avec une pièce de monnaie…on a vu … parfois on gagne, parfois on ne gagne pas…mais en moyenne…on gagne combien ? »* doivent s'inscrire dans une logique de construction de nouveaux critères de validation en classe des mathématiques.

Bref, il nous semble improbable l'élaboration autonome du principe de maximisation de l'espérance de la part des élèves dans une classe ordinaire des mathématiques. Les rôles de chacune des parties étant assez bien définis, nous obligent à prévoir donc l'intervention de l'enseignant pour encourager et/ou valider les débats et les propositions allant dans la direction d'une argumentation dans un cadre indéterministe.

Quatrième étape : les tendances

Avec une pièce de monnaie

Nous cherchons à associer le calcul de probabilités à la stabilisation de fréquences. Pour ce problème le premier sera justifié comme un moyen économique d'anticiper le deuxième. Nous prévoyons des aller-retour entre l'un et l'autre afin de renforcer leur lien, et cela pour tous les calculs de probabilités concernés.

Une fois que les élèves aient commencé à s'intéresser à la moyenne de points gagnés par nombre de pièces, ils pourraient calculer celle d'une pièce de monnaie. Facile à évaluer, cette tendance ne remettrait pas en question les connaissances des élèves : ils sauront qu'en moyenne, le gain sera de 5 points.

La terminologie spécifique (espérance, probabilité) serait travaillée lors de la dernière étape. Il nous semble plus important que les élèves utilisent des termes variés, en particulier pour se référer à la probabilité, de même que les évaluations numériques se fassent en termes de pourcentage ; type de représentation qui s'avère plus adaptée à la sémantique de la probabilité dans ce problème. Vu notre intention d'institutionnaliser la notion fréquentiste, cette terminologie variée devrait être récupérée par l'enseignant afin de décrire l'attente à long terme. Rappelons que les deux notions de la probabilité ne peuvent se différencier que sur le registre langagier et en employant des termes synonymes.

Avec deux pièces de monnaie

Pour déterminer la moyenne de points à gagner pour deux pièces de monnaie, nous nous attendons à une erreur courante, celle de considérer trois possibilités équiprobables : (pile, pile) ; (pile, face) ; (face, pile). L'évaluation numérique de la probabilité étant d'un tiers, la moyenne de points gagnés serait de 6,67 points.

Nous proposons de tirer profit de cette erreur pour discuter en classe le principe de la loi faible des grands nombres. P proposerait de tester le calcul de la probabilité *a priori* par des tirages de deux pièces de monnaie. Mais l'hypothèse fausse (« la proportion de (pile, pile) est 1/3 ») est très proche de la hypothèse vraie 1/4, et même avec un nombre important de données la convergence de l'échantillon vers ¼ pourrait ne pas s'éloigner suffisamment de 1/3 (voir Tableau 55).

Distances entre hypothèses vraies et fausses

	2 piles	1 pile 1 face	2 faces
Hypothèse Vraie	¼	½	¼
Hypothèse Fausse	1/3	1/3	1/3
Distance	1/12	1/6	1/12

Tableau 55

Nous proposons donc de la confirmer d'une manière indirecte, en le faisant sur le couple (1 pile, 1 face). Ces deux hypothèses sont deux fois plus distantes que les autres couples. P proposerait donc de confirmer si l'événement (1 face, 1 pile) se produit à long terme avec une fréquence de 33 % des fois. Pour cela, P solliciterait dix lancers par élève, en supposant vingt élèves dans la séance, la taille de l'échantillon serait de 200 données.

Nous n'avons pas l'intention de simplement comparer la proportion finale de l'échantillon avec celle déterminée *a priori* par les élèves. Nous souhaitons que la convergence en fonction de la taille de l'échantillon soit aussi analysée. Pour cela les données seront introduites progressivement, d'abord la première vingtaine, puis ajouter une autre vingtaine, etc., afin de mettre en évidence la convergence.

Pour cela, les élèves dessineront une table, en écrivant 1 s'ils obtiennent le résultat en question, 0 pour les autres événements. A la fin des tirages, chacun aura une table avec ses lancers (Tableau 56) :

Tableau pour l'étude de la loi faible des grands nombres. Vingt lancers par élève

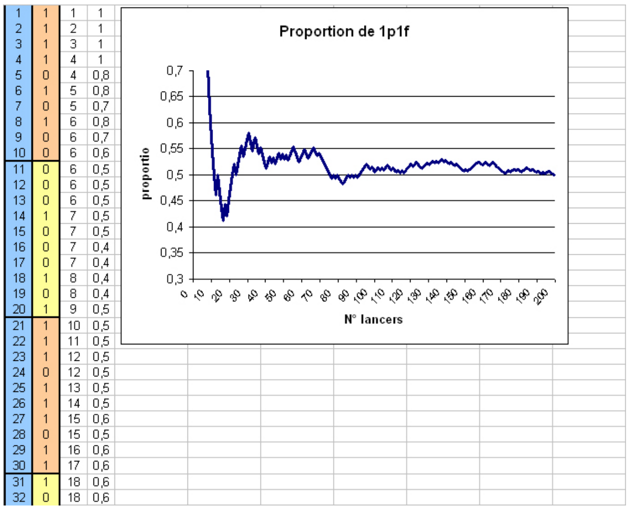
Tirage	1	2	3	4	5	6	7	8	9	10
Résultats	1	1	1	0	1	0	1	0	0	0

Tableau 56

S'il y a des ordinateurs dans la salle, chaque élève enregistrerait sur un fichier ses données. P réunirait les données dans un même fichier en les stockant sur une clé usb. Ce fichier aurait été préalablement préparé pour la représentation graphique de la convergence de la proportion de (1 pile, 1 face). L'image serait projetée sur un écran (Figure 16). Si la salle ne dispose pas d'ordinateurs, le travail d'enregistrement se ferait sur des ordinateurs portables.

Lors de la représentation graphique, il conviendrait de choisir une échelle appropriée pour mettre en évidence la convergence vers ½ et non pas 1/3.

Capture d'écran de l'image à projeter dans la salle



1	1	1	1
2	1	2	1
3	1	3	1
4	1	4	1
5	0	4	0,8
6	1	5	0,8
7	0	5	0,7
8	1	6	0,8
9	0	6	0,7
10	0	6	0,6
11	0	6	0,5
12	0	6	0,5
13	0	6	0,5
14	1	7	0,5
15	0	7	0,5
16	0	7	0,4
17	0	7	0,4
18	1	8	0,4
19	0	8	0,4
20	1	9	0,5
21	1	10	0,5
22	1	11	0,5
23	1	12	0,5
24	0	12	0,5
25	1	13	0,5
26	1	14	0,5
27	1	15	0,6
28	0	15	0,5
29	1	16	0,6
30	1	17	0,6
31	1	18	0,6
32	0	18	0,6

Figure 16

Cette étape d'expérimentation et de représentation graphique devrait permettre à la classe de se rappeler des notions apprises au lycée sur le calcul de probabilités, principalement celle de la décomposition en cas élémentaires équiprobables. Les techniques (axiomes de Kolmogorov) de calcul de la probabilité seraient alors rappelées. Les probabilités pour chaque issue possible seraient recalculées, de la même manière que l'espérance mathématique des points à gagner pour deux pièces de monnaie.

Si nécessaire, l'expérimentation pourrait se réaliser pour le choix de trois et quatre pièces. Une fois calculées toutes les tendances, la classe seraient en conditions de déterminer la stratégie gagnante et de la justifier.

Cinquième étape : l'institutionnalisation

A la fin de la séance des aspects sémantiques et opératoires devraient être institutionnalisés.

Le signifié de la probabilité

La probabilité se réfère ici à la tendance d'apparition à long terme d'un certain phénomène. Il faut remarquer l'importance du nombre des données dans une approximation par un échantillon.

Le signifié de la probabilité serait institutionnalisé à partir de deux sources :

- Le contexte de la situation, principalement par l'identification de la nature de l'objet, dans ce cas-ci, la série infinie et ses principales caractéristiques (indépendance des événements élémentaires, mêmes conditions de reproduction et convergence lorsque la taille tend vers l'infini.
- D'autre part les échanges langagiers que l'enseignant aurait pu retenir. Ceux entre les élèves seraient analysés par leur correspondance avec la nature de l'objet de l'item précédent.

La représentation symbolique

Symboliquement, la tendance d'apparition est représentée par P(A). A, dans ce cas, la série infinie, disposant des caractéristiques ci-dessus énumérées.

Le calcul

Pour un calcul *a priori* : rappel de formule de Laplace, décomposition en cas élémentaires et axiomes. Pour une estimation par échantillon, mentionner l'importance de sa taille.

L'espérance mathématique

Présenter l'expression mathématique du calcul de l'espérance mathématique, sans trop de formalismes.

Le critère décisionnel

Retenir le critère qui a permis de décider.

Analyse a posteriori

La salle et son équipement

La Figure 17 est une représentation de la distribution des vingt deux élèves et du matériel informatique utilisé pour la séance. La salle disposait d'un ordinateur (1) sur le bureau de l'enseignant relié à un tableau tactile (2) et d'un vidéo-projecteur (3). Nous avons placé une camera (4) au fond de la salle et deux dictaphones (5 et 6) sur les tables des élèves. Finalement nous avions à disposition deux ordinateurs portables (7).

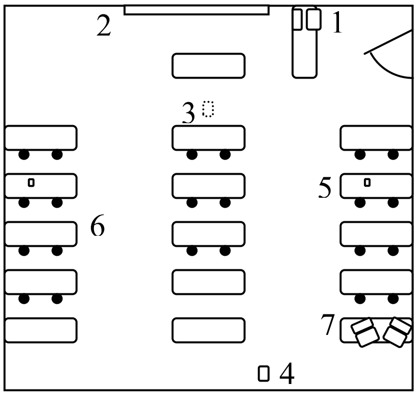
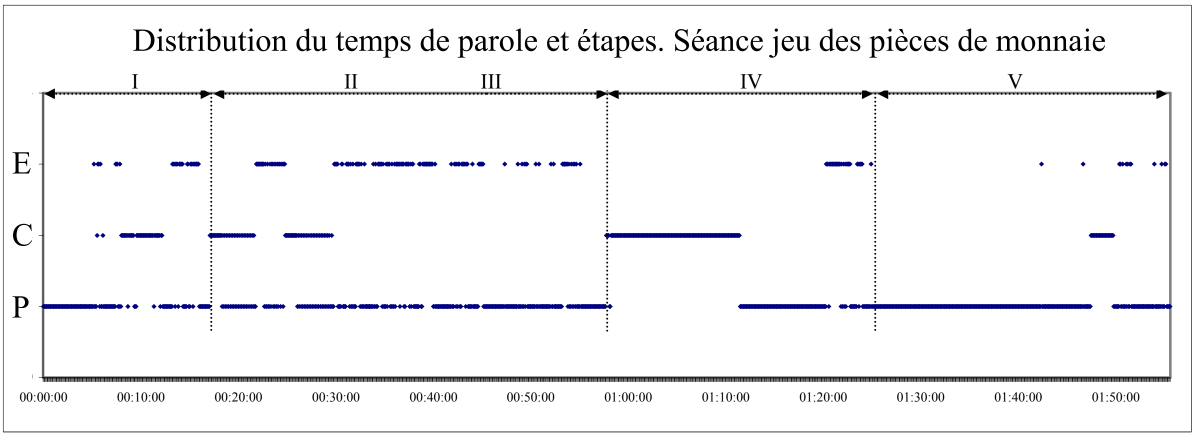


Figure 17

Le Graphique 37 est une représentation de la séance qui s'est déroulée le 13 novembre 2006 où nous avons expérimenté cette première situation-problème. Pour ce graphique nous avons pris en compte les échanges de parole en fonction de trois groupes. Notre intérêt pour ce genre d'échanges s'explique par les possibilités d'expression qu'offre le registre langagier aux interprétations de la probabilité et les prises de décision. Les groupes retenus sont :

- Groupe P : L'émetteur du message est l'enseignant et il s'adresse à toute la classe.
- Groupe E : L'émetteur du message est un élève et il s'adresse à toute la classe.
- Groupe C : L'émetteur (enseignant ou élève) s'adresse à un membre de la classe ou à un groupe. Généralement, ce groupe étant constitué par des élèves voisins travaillant ensemble pour résoudre le problème.



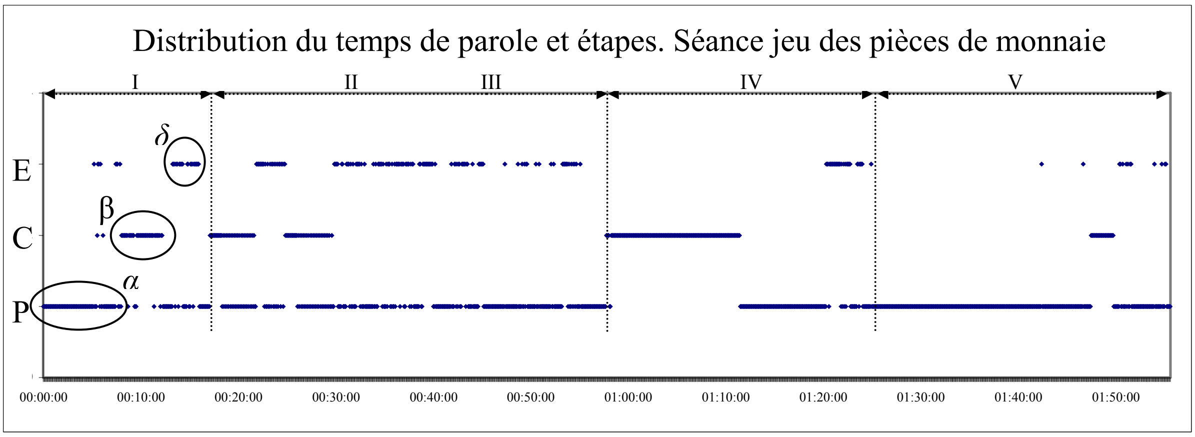
Graphique 37

Pour l'élaboration du graphique nous avons pris en compte le critère suivant : A chaque fois qu'un membre d'un de ces groupes émettait un message, nous notions un point dans l'échelle chronologique. A chaque intervalle de cinq secondes de parole nous ajoutions un autre point. En annexe nous avons placé le décryptage des échanges enregistrés.

Ce graphique nous renseigne de la distribution de parole au long de la séance. Sur ce graphique nous avons repéré les quatre étapes de notre analyse *a priori* (exploration, assemblement de données, validation du critère, les tendances et institutionnalisation). Nous nous servirons de cette représentation pour notre analyse *a posteriori*, nous commençons par les quatre étapes prévues dans notre analyse *a priori*.

Exploration du problème

Cette étape se déroule pendant les quinze premières minutes de la séance (Région I du Graphique 37).L'enseignant présente le problème aux élèves (voir α dans le Graphique 38). Suite à quelques questions concernant les consignes du jeu, les élèves commencent à s'approprier le problème et lancent leurs pièces de monnaie (β). Dans cette phase de découverte, les échanges se produisent plutôt entre paires ou au sein de chaque groupe de travail (deux élèves, occasionnellement quatre).



Graphique 38

Après avoir évacué les doutes concernant les règles du jeu, l'enseignant interroge les élèves en cherchant à installer un débat autour de leurs premières impressions. Pendant ce débat (δ), l'enseignant cherche à approuver la participation des élèves en reproduisant leurs opinions. Il anime et modère les échanges, en évitant toujours d'interpeller sur les raisons des différentes propositions, probablement pour ne pas inhiber les élèves de s'exprimer. Des interventions, nous retenons plusieurs éléments :

- *Les variables*. Les élèves reconnaissent les variables du problème, ils parlent du bénéfice de jouer avec quatre pièces par les points que cela remporte mais aussi de la rareté de réussir, et vice-versa : de la facilité d'obtenir des points lorsque l'on joue avec une pièce mais du faible profit que cela ramène.
- *Le signifié de la probabilité*. Les élèves utilisent les mots *chance* et *rare* pour se référer à la probabilité. Pour le premier mot, il nous est impossible d'assurer la nature de l'objet auquel ils font référence (série ou épreuve générique). Pour le moment personne n'emploie ces termes pour évoquer la probabilité de gagner le jeu (probabilité bayésienne, de nature exclusivement qualitative, Elèves en **gras**, enseignant en *italique*):

13 **(…) que plus on prend de pièces moins on a de chances d'avoir que des piles**
(…)
Si on joue avec quatre pièces moins de chances de gagner
En un lancer
Moins il y a des pièces plus on a des chances de gagner
(…)
Bah si on joue avec une pièce on a plus de chances de tomber sur pile si on a beaucoup de pièces on a moins de chances de tomber sur de pas tomber sur face plutôt
Ouais
Bah le nombre le nombre de points récoltés avec un grand nombre de pièces est grand mais plus rare aussi

- *Des stratégies pour gagner*. Aucune stratégie n'émerge encore. Sous l'hypothèse implicite d'une régularité ou stabilité du jeu, les élèves transposent leur vécu et le généralisent. Ainsi description des lancers effectués et inférence se confondent. Ces inférences ne se formalisent pas, la classe est en train de découvrir les deux variables en question.
- *L'opposition entre variables.* Les points gagnés et la fréquence d'apparition sont perçus comme s'opposant l'un à l'autre, aux quarante points s'oppose leur rareté ; à la facilité, la faible ponctuation accumulée. Cette opposition fait que les variables se

bloquent mutuellement. Lorsque les élèves considèrent l'une, ils négligent l'autre, ainsi la perception du problème est encore uni-variée.

- *Evaluation comparative*. Les élèves ne formulent pas d'hypothèses, cela est accompagné de l'utilisation de quantificateurs comparatifs (plus rare, moins souvent). Ainsi les fréquences d'apparition sont relatives au couple d'événements comparés (quarante apparaît plus rare que dix), une sorte d'évaluation comparative (par couple d'événements) et non pas par une échelle.

Rassemblement des données

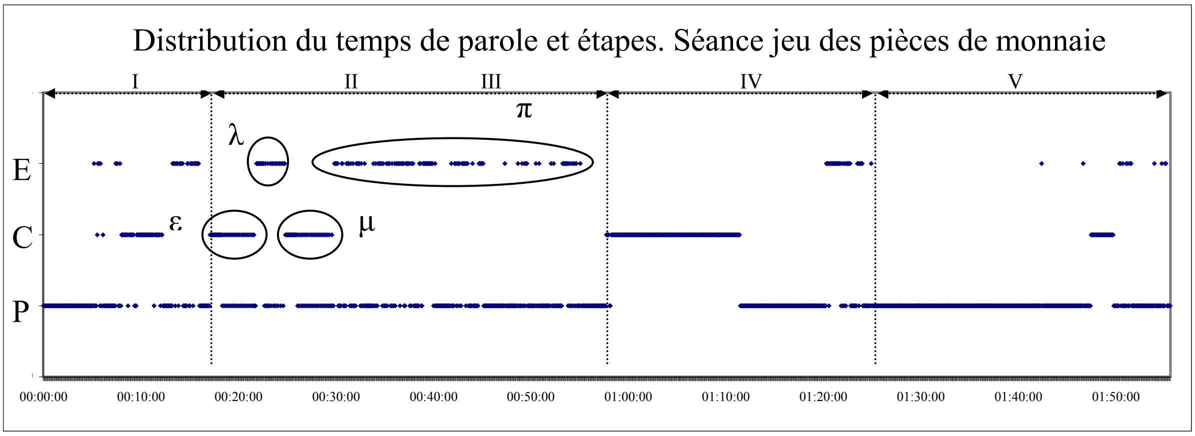
Première série de données

Cette deuxième étape commence à la seizième minute, lorsque l'enseignant intervient pour proposer la production et le rassemblement de données. Au moment où l'enseignant demande d'effectuer les lancers, les élèves se trouvaient encore dans une position dichotomique, de choisir une variable ou l'autre.

En plus, la demande de l'enseignant de nouveaux lancers n'a pas été accompagnée des objectifs que cette nouvelle tâche visait. Ainsi, une sorte de rupture se produit dans la séance. Jusqu'à ce moment les élèves étaient en train de participer, de s'exprimer, de raisonner. Soudainement, l'enseignant leur demande de réaliser une tâche dont ils ne connaissent pas les raisons.

Probablement que cette demande si rapide puisse s'expliquer par le souci de l'enseignant à respecter le planning prévu pour la séance, planning qui coïncide en général avec les étapes de notre analyse *a priori* (exploration, rassemblement de données, etc.). En effet, lors de la préparation de cette séance nous avons discuté avec lui d'un possible déroulement. Pour mieux l'analyser ensemble nous lui avions fait parvenir une copie d'une analyse *a priori* assez proche de celle présentée dans cette synthèse. Ainsi nos étapes, ayant accompagné tous nos entretiens, sont devenues celles de l'enseignant.

Toutefois, les élèves effectuent les quatre lancers demandés (ε, Graphique 39), le premier avec une pièce, le deuxième avec deux pièces, etc. A la dix-huitième minute, l'enseignant commence à noter les résultats sur le tableau tactile.



Graphique 39

Première stratégie. La contingence des données

Une fois le tableau de données complété, l'enseignant demande de l'analyser, de cette manière la classe reprend le débat (λ). Les élèves, ayant précédemment souligné en δ l'importance des variables *rareté* et *points gagnés,* se voient confortés par les données du tableau, et ils le font savoir, le débat reprend donc par les deux variables du problème.

De la même manière qu'en δ dont un élève a parlé de *«* **ce qui arrive en général** *»*, un autre en λ glisse la mesure de la rareté de deux événements : *«* **Une chance sur deux et une chance sur quatre** *»*. En s'appuyant sur des hypothèses d'équipossibilité, cet élève quitte le stade d'évaluation comparative et assigne des valeurs à la fréquence d'apparition sous la forme d'une proportion : c'est la première manifestation d'une évaluation numérique de la probabilité. Néanmoins, cette proposition n'a été retenue ni par l'enseignant ni par le reste des élèves.

22 **Bah qu'il y a un seule qu'a réussi à avoir quarante le reste**
Que les dix ça va vite
Que les dix ça va mais le reste
Une chance sur deux et une chance sur quatre
Comme on a dit tout à l'heure plus il y a de pièces
Attends comment tu dis
Que plus il y a de pièces moins on a de chances de réussir
Pas la grosse chance
Il faut que le jeu soit équitable
On continue
C'est pas équitable du tout pas du tout il faut voir les points
Il suffit de faire il suffit de faire la somme de chaque colonne quoi et voir le nombre de points qu'on a eu par colonne quoi
Bon faisons çà rapidement il nous suggère de faire la somme donc la somme pour quatre c'est facile ici quarante donc ici la somme c'est quarante je ferai un deux trois quatre donc la somme c'est quarante

Il pourrait paraître étrange que l'enseignant choisisse de suivre la voie proposée par le dernier élève, faire la somme par colonne pour trouver la stratégie gagnante. Néanmoins cela s'expliquerait pour plusieurs raisons :

Premièrement, pour la première fois une stratégie s'esquisse. Jusqu'à là, personne n'avait proposé un critère, la classe doit donc l'analyser.

Deuxièmement, cette stratégie avait été anticipée lors de nos entretiens, et nous en avions proposé une suite, refaire l'expérience et comparer les résultats des deux séries. Avec des échantillons de taille vingt deux, les sommes à comparer seraient très probablement différentes. Ainsi, la stratégie de la somme par colonne serait facilement réfutée.

Le fait d'avoir anticipé cette stratégie et de disposer des outils pour en agir facilite le choix de l'enseignant. En entreprenant une voie inscrite dans le planning, il vera réduit le stress de la gestion du déroulement.

Troisièmement, les notions utilisées dans les arguments de réfutation de cette stratégie sont d'utilité pour traiter la stabilisation de fréquences (fluctuation de l'échantillonnage, hasard, etc.). Le temps donc de sa réfutation est, pour l'enseignant, un temps d'investissement.

Etude de la première stratégie

Les élèves suivent l'indication de l'enseignant, ils procèdent donc en calculant les totaux par colonne (λ). Les sommes des points correspondantes à une et trois pièces sont supérieures à celles de deux et quatre. Les opinions des élèves sont divergentes, pour un élève les sommes ne sont pas relevantes, il déclare que tous les choix sont équivalents (*«* **Pour moi c'est pareil** *»*). Il développera en μ ses arguments, nous y reviendrons. D'autres reconnaissent le caractère aléatoire de la série de vingt deux lancers et refusent de baser leur stratégie à partir des sommes effectuées (*«* **Mais non parce qu'il faudra faire plusieurs séries comme ça ça peut pas être un cas général ça** *»***)**. L'enseignant avance donc dans cette dernière direction :

24 *Ici on n'a fait on n'a fait que vingt deux vous avez fait chacun un tirage est si on fait on refait la même expérience est-ce qu'on va obtenir la même chose*

Non non

(…)

Est-ce qu'on va obtenir la même chose est-ce qu'on va obtenir la même chose est-ce qu'on va obtenir quelque chose de contraire

Bah non on ne sait pas

(...)

On ne peut pas parce qu'il y a du hasard
C'est du hasard donc on ne peut pas savoir
Mais il y a quand même
Mais c'est une approximation
Il y a quand même une probabilité
Attendez attendez
Il y a quand même une probabilité que l'un reste dans les alentours de cent cinquante bah pas exactement cent cinquante mais
Bah on fait l'expérience si vous voulez

Confrontation de paradigmes. Déterminisme versus indéterminisme

Dans les échanges des élèves du tableau ci-dessus evoquent deux paradigmes logiques bien différents :

- Un élève qui reste attaché au déterminisme, avec ses valeurs logiques associées : il ne sait pas ce qu'ils obtiendraient s'ils répètent l'expérience. Sa position bascule entre deux valeurs, savoir (vrai faux) ou ne pas savoir. S'il ne sait pas, il ne peut pas se prononcer.
- Un autre élève qui aborde le sujet à partir d'une approche indéterministe. Contrairement à la position dichotomique de son collègue, il place son opinion entre les valeurs logiques vrai et faux. Il n'a pas la certitude, mais (*«* **quand même** *»*) l'impression qu'une nouvelle expérimentation donnerait des valeurs proches. Pour la première fois le mot probabilité est employé et il nous semble que, par sa place dans ces échanges, le sens de la probabilité représente une mesure de certitude. Ces deux approches seront décisives pour répondre à la question du problème. De la position que prennent les élèves dépendra la suite des débats.

Suite à la proposition de l'enseignant (refaire l'expérience), les élèves effectuent des nouveaux lancers, quatre par élève dont un pour chaque choix possible (μ). A la différence des précédents (ε) dont il y a eu une rupture dans la séance, ces lancers se sont effectués avec une intention déclarée : tester si les sommes d'une nouvelle série se rapprochent aux sommes des premières.

A la minute vingt huit de la séance (π), les élèves calculent les sommes par colonne de la nouvelle série. Ces résultats ont inversés l'ordre des choix des pièces de monnaie :

Bah auprès le premier tableau celui qui gagne le plus c'est une seule pièce
Et trois pièces
Et trois pièces aussi tandis que dans le deuxième tableau c'est plus le cas c'est différent c'est quatre c'est deux et quatre

Avec ces nouvelles données, la classe perd la seule stratégie dont elle disposait jusqu'à ce moment, c'est-à-dire, choisir les pièces qui ont sommé le plus de points. Cela laisse la classe quelque part perplexe, les données n'ayant rien de significatif à dire :

Il y a du hasard

les idées que vous aviez sur le jeu alors vos réactions on entend c'est du hasard qu'est-ce que tu veux dire par hasard

30 **Qu'on ne sait pas du tout comment va sortir pour une pièce pour deux pièces pour trois pièces on ne peut pas savoir**

Face à ce manque de stratégie, un élève, celui qui avait proposé l'inexistence de stratégie (*«* **Pour moi c'est pareil** *»*) revient et développe ses arguments basés sur une sorte de géométrie du jeu. Son approche est différente, vue que les données ne lui ont rien apporté de significatif, il fonde ses raisons sur la compensation qui se produit entre la rareté et les points gagnés.

Cette nouvelle stratégie est à partir de ce point de vue, plus riche que la précédente : elle se construit sur une analyse de la structure du jeu, en non pas directement sur des évidences factuelles si faibles en nombre. De plus, cette stratégie récupère les variables que le groupe avait identifiées auparavant.

La place que donne cet élève aux deux séries de vingt tirages reste une question ouverte, nous ignorons si pour lui l'écart entre les séries est dû au faible nombre de données ou si par contre et d'une manière sélective, il ignore cette information que le contredit. Toutefois, quelques élèves insistent sur la stratégie de choisir une pièce. Pour la défendre, ils trouvent des exemples en guisse d'illustrations, dans le premier et le deuxième tableau. Ils signalent des lignes dans ces deux tableaux dont les points cumulés avec une pièce dépassent ceux de quatre pièces. L'enseignant, de sa part, signale sur les mêmes tableaux des contre exemples.

Entre les illustrations des uns et les réfutations des autres, un autre élève essaie d'introduire son approche, il insiste et il le fait à trois reprises: *«***une chance sur deux** *»*. C'est le même qu'en δ a glissé *«* **une chance sur deux et une chance sur quatre** *»* mais il n'accompagne toujours pas ses déclarations par des explications ni leurs applications. La classe, par ses propres moyens, semble ne pas être en mesure de s'apercevoir de la signification de ces propositions.

De l'évaluation comparative vers la numérique

Au même temps, (minute trente quatre) quelques élèves insistent sur la stratégie d'une pièce de monnaie. Etant l'argument basé sur les données déjà réfuté par l'enseignant, ils en cherchent un autre. Ainsi un élève, en regardant le tableau, revient sur la rareté des événements, un autre élève étaie un autre point de vue :

Voilà par exemple le vingt c'est impossible c'est impossible d'avoir trois vingt quasiment d'affilé
C'est pas possible d'avoir trois vingt d'affilé
C'est beaucoup plus dur quoi
Si on regarde le tableau en bas on a deux fois quarante de suite
Mais par rapport au nombre de faces il va avoir une chance sur deux de gagner d'avoir dix
Ouais
Ouais

De cette manière se produit le passage dès l'évaluation comparative vers la numérique dans les échanges de ces deux élèves, le premier, encore dans la position d'évaluation, comparative s'adresse à la classe en termes de *«* **impossible** *», «* **dur** *»* ; l'autre élève s'abstrait des contingences et modélise la situation : *«* **Mais par rapport au nombre de faces il va avoir une chance sur deux de gagner d'avoir dix** *»*. Notons que la proposition énoncée à partir du stage d'évaluation comparative est réfutée (*«***Si on regarde le tableau en bas on a deux fois quarante de suite** *»*), néanmoins cette réfutation perd son intérêt à partir de l'énonciation de l'évaluation numérique (*«* **ouais** *», «* **ouais** *»*).

D'ailleurs, les caractéristiques de ce jeu permettent aux élèves de modéliser facilement une des stratégies qu'ils considèrent comme étant la gagnante (elle en est une quand même). Pour une pièce de monnaie, les événements (gagner ou ne pas gagner) sont équiprobables, les élèves n'ont pas de difficultés à formuler leurs propositions en termes de moyenne. Bien que leurs idées s'expriment de manière banale, elles sont sémantiquement proches de l'espérance mathématique :

En dix lancers normalement on arrive à cinquante
Ouais en dix lancers
En dix lancers
On a une chance sur deux de tomber sur pile en dix lancers normalement on devrait arriver à cinquante

Dans le discours de ces élèves le regard à long terme et la compensation qui se produit entre les données sont sous-entendu. Lorsqu'un autre élève, placé a partir de l'approche déterministe essaie de réfuter leurs idées, ils se voient obligés de le préciser :

Sur dix lancers tu peux toujours tomber sur face
Oui en moyenne en moyenne
En moyenne
En moyenne
C'est une moyenne quoi

Afin de consolider quelques notions et ainsi avancer, l'enseignant intervient, il valide la notion de "chance sur deux" pour toute la classe et en cherchant à soutenir la nouvelle approche (indéterministe) il demande au groupe d'élèves de préciser, pour tous, le signifié de moyenne. Néanmoins, les positions sont encore fragiles, les élèves qui ont défendu cette nouvelle approche ne se sentent pas surs de leurs affirmations et il déclinent l'invitation de l'enseignant (« **je me suis mal exprimé** »).

Les positions se confrontent, les élèves placés à partir d'un point de vue déterministe discutent les arguments de ceux dans l'indéterminisme. Les signifiés des expressions sont différentes pour chaque approche, « **En dix lancers** (…) **on arrive à cinquante** » n'est pas interprété de la même manière pour les uns et pour les autres. Pour les premiers cela signifie *une* série de dix lancers ; pour les deuxièmes l'intention est toute autre, ils parlent en termes de moyenne.

Lors de ce débat de dix minutes, la conviction des "indéterministes" se détériore lorsque par exemple ils ne peuvent pas assurer d'accumulation de cinquante points pour une série de dix lancers. Ils se replient sur les propositions que ne sont pas remises en cause par les "déterministes". Ce sont des notions relatives au calcul de la probabilité, («**une chance sur quatre** (…) **une chance sur seize** »). Ces notions ne sont contestées par personne, d'une part parce qu'elles ont été validées institutionnellement, au lycée et dans la séance par l'enseignant, d'autre par parce qu'elles ne concernent pas des inférences indéterministes (déduction, voir *Chapitre I*).

Pendant ces dix minutes, les deux positions développent leurs arguments et consolident leurs positions, principalement les déterministes. Pour eux, l'argument tourne sur l'impossibilité de se prononcer lorsqu'on n'est pas certain et même s'ils admettent une possible modélisation en termes d'attente à long terme, cela ne leur satisfait pas :

Ouais la moyenne on peut la faire mais c'est par autant que même si par exemple sur la pièce de jeux on a une bonne moyenne par exemple par rapport à la chance c'est par pour autant
Oui mais c'est théoriquement ça
(…)
Oui mais là tu peux pas te fier sur laquelle tu ne peux pas te fier pour faire tes lancers

Validation du critère

A la minute quarante cinq, l'enseignant intervient et débloque la situation (étape III de notre analyse *a priori*). Il indique le critère à suivre en présentant le principe de maximisation de l'espérance comme argument décisionnel. L'enseignant se voit donc obligé à résoudre le conflit de paradigmes et il le fait en validant les raisonnements du groupe indéterministe, groupe qui avait avancé jusqu'à même s'intéresser aux pourcentages lors d'une "infinité de tirages" pour finalement se plier et se limiter aux algorithmes de calcul, seules propositions n'ayant pas été réfutées par les déterministes.

Cette validation par l'enseignant, est suivie d'une série de tâches pour les élèves : « *on va s'intéresser (...) pour chacune des possibilités (...) on va regarder combien (...) en moyenne on obtient de points par tirage et puis en faisant çà on aura peut-être à se prononcer pour une stratégie la meilleure (...) »*.

Avant que les élèves déterminent les moyennes respectives, l'enseignant les interroge sur le signifié de ce terme. Les réponses sont mitigées, et elles correspondent en fait à des questions qu'il n'a pas posées. L'enseignant décide donc d'approcher le concept de moyenne à partir des deux séries de tirages effectuées. Les élèves calculent les moyennes des deux séries pour une pièce de monnaie et l'enseignant élabore les conclusions.

Les rôles ont changé depuis que l'enseignant est intervenu pour valider l'approche indéterministe. Avant, il laissait assez libre la gestion de la séance (plus de dix minutes). Avec la validation, qui se produit par l'impossibilité des élèves de résoudre un tel conflit d'approches, il y a une accélération de tâches dont les rôles sont assez polarisés. Les élèves réalisent des activités plutôt simples (calculatoires) et l'enseignant de sa part élabore les conclusions et réassigne des nouvelles tâches.

Par exemple, lorsque les élèves ont calculé les moyennes des deux séries, l'enseignant reprend ces calculs, décide de rassembler les deux séries en une de quarante quatre lancers puis il renvoie aux élèves la tâche de calculer la nouvelle moyenne correspondant à la réunion des données. Finalement, il élabore les conclusions en expliquant la convergence du gain moyen vers la valeur cinq que quelques élèves avaient postulé auparavant.

Le gain moyen pour deux pièces

Pour une pièce de monnaie, les deux événements possibles (gagner ou ne pas gagner) sont équiprobables, et le gain moyen est assez rapidement compris par les élèves. Néanmoins, ils ne réagissent pas si facilement lorsque l'enseignant demande d'appliquer le même procédé pour calculer le gain moyen pour deux pièces de monnaies. Un seul déclare la valeur en question (« **cinq** »). Mais le reste de la classe n'est pas en mesure de comprendre la démarche effectuée par cet élève (« **Tu ne peux pas le sortir comme çà sans rien expliquer** »).

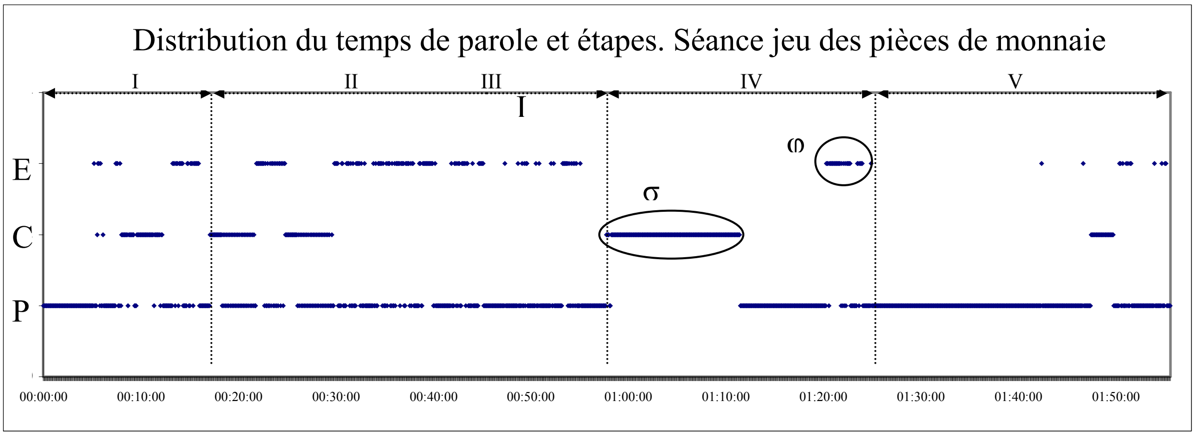
Si pour une pièce de monnaie la place de la probabilité pourrait passer inaperçue dû à l'équivalence entre moyenne pondérée et moyenne arithmétique, ce n'est plus le cas pour le choix des deux pièces. L'élève qu'a déclaré cinq comme moyenne se voit obligé à décliner son explication en précisant pas à pas sa méthode, les cas favorables et possibles puis la valeur de la probabilité et finalement le gain moyen.

L'écart entre cet élève et le reste de la classe est évident. La clarté de son énonciation et le formalisme des termes employés laissant perplexe la classe qui ne fait qu'effectuer les tâches calculatoires que l'enseignant leur donne lors de la révision du procédé. L'explication qui fait l'élève du calcul du gain moyen pour deux pièces de monnaies est quasiment acceptée par les élèves, même si quelques-uns semblent l'avoir fait juste par l'impression provoquée par la propreté de l'explication.

Les tendances

Tout de suite (minute cinquante quatre, σ en Graphique 40), l'enseignant interroge la classe en demandant si le procédé de calcul de la probabilité de gagner avec deux pièces a été bien compris, et il le fait avec une certaine insistance. En fait, dans notre planning nous avions déjà prévu l'étude de la stabilisation de fréquences (*Chapitre III. Analyse* a priori*. Quatrième étape : les tendances*). Une possibilité pour l'introduire consistait à profiter d'une erreur usuelle dans le calcul de la probabilité de gagner des points avec deux pièces. Sous le prétexte de vérifier ce calcul (supposé mal fait), les élèves effectueraient l'étude de la convergence.

Néanmoins, l'erreur attendue ne se pas produite, le procédé indiqué par l'élève étant le correct, l'enseignant décide donc d'arriver à l'étude des tendances par une voie alternative. Il s'appuie sur les doutes de quelques élèves. C'est en testant la méthode correcte et ne pas l'incorrecte, comme il était prévu, que l'enseignant introduit dans la classe l'étude de la convergence de la série.



Graphique 40

L'intention d'introduire l'étude de la stabilisation de fréquences explique donc l'insistance de l'enseignant pour trouver des élèves non convaincus par l'explication du calcul de la probabilité. Ayant trouvé des élèves hésitants, l'enseignant demande d'effectuer des tirages pour confirmer le calcul. La nouvelle consigne consiste à lancers dix fois deux pièces de monnaie par élève, ils noteraient 1 (un) pour le couple (pile, pile) et 0 (zéro) pour les autres cas.

Suite aux lancers effectués, les élèves enregistrent les données sur des fichiers localisés sur les ordinateurs portables (Figure 18). Après avoir récupéré les données des deux ordinateurs à l'aide d'une clé usb, l'enseignant les transfère vers son ordinateur relié au vidéo projecteur. Cette nouvelle production de données se déroule dès la minute cinquante huit jusqu'à la soixante et onze (voir J en Graphique 40). Ce temps inclut une pause pour les élèves.

Fichier pour l'enregistrement de tirages des élèves

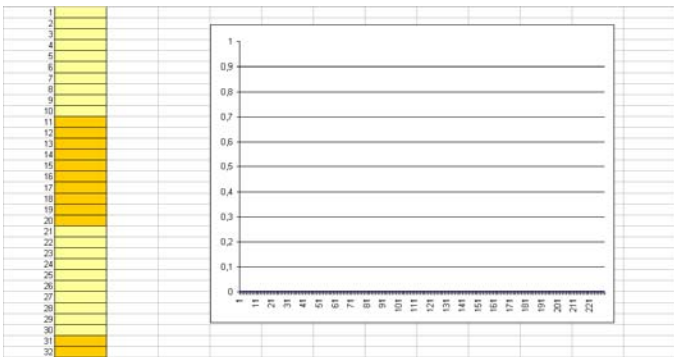


Figure 18

Pendant la période qui suit le rassemblement des données et jusqu'à pratiquement la fin de la séance, la convivialité entre professeur et élèves s'est vue érodée. Probablement à cause de la fatigue provoquée par cette activité longue. De plus, notre présence peut avoir contribué à augmenter le stress tant des élèves que de l'enseignant.

En se servant du vidéo-projecteur l'enseignant montre à la classe le graphique de l'évolution de la proportion d'apparition du couple (pile, pile) (Figure 19). Il explique les différents éléments de la feuille de calcul. En particulier, que la dernière colonne correspond à « *l'évolution de ce calcul de pourcentage qu'on soit bien d'accord ».* En suite il cherche à que les élèves tirent des conclusions en comparant le graphique avec le calcul de la probabilité récemment effectué.

Image projetée de la proportion de (pile, pile)

1	0	0	0
2	1	1	0,5
3	0	1	0,3333
4	0	1	0,25
5	1	2	0,4
6	0	2	0,3333
7	0	2	0,2857
8	1	3	0,375
9	0	3	0,3333
10	1	4	0,4
11	0	4	0,3636
12	0	4	0,3333
13	1	5	0,3846
14	1	6	0,4286
15	1	7	0,4667
16	0	7	0,4375
17	0	7	0,4118
18	1	8	0,4444
19	0	8	0,4211
20	0	8	0,4
21	1	9	0,4286
22	0	9	0,4091
23	0	9	0,3913
24	0	9	0,375
25	1	10	0,4
26	1	11	0,4231
27	1	12	0,4444
28	0	12	0,4286
29	1	13	0,4483
30	1	14	0,4667
31	0	14	0,4516

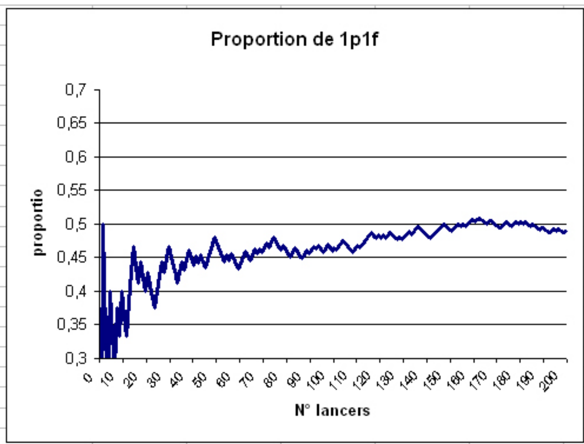


Figure 19

Des deux dimensions à relier (numérique et sémantique), les élèves développent plus facilement la numérique, les raisons semblent être multiples, parmi elles :

Les possibles approches sur le graphique sont nombreuses et ouvertes, à la différence de la dimension numérique que peut se réduire à un algorithme. De plus cette dernière a été déjà institutionnalisée. Les élèves disposent donc de plus d'outils pour y réagir.

- Le graphique est une nouveauté, y compris ses éléments. Cela augmente le risque de se tromper. Par exemple, un élève essaie de commencer par l'interprétation d'une cellule du graphique mais sa terminologie est inappropriée (utilisation du terme probabilité au lieu de pourcentage) et l'enseignant corrige tout de suite.

Finalement, les élèves abordent le lien à partir de la dimension numérique (φ, Graphique 40). La terminologie employée n'est pas celle traditionnellement institutionnalisée (cas favorable et possibles), ils s'expriment en termes banalisés : **« une chance sur quatre ».** Ce vocabulaire non seulement n'est pas rejeté par l'enseignant mais même accepté et souvent utilisé. Le terme *chance*, bien qu'imprécis, renferme dans son ambiguïté le signifié fréquentiste (Robert & Collins, 2001) ; il n'y aurait pas donc une erreur dans son emploi mais plutôt un manque de précision. Ce manque de précision serait compensé par des éléments du contexte qui contribueraient à donner du sens au terme. Il est arrivé dans cette séance que les sens attribués à certains termes ont différé d'élève à élève, ainsi l'émetteur du message s'est vu obligé à le préciser. A ce moment de la séance, l'ambiguïté du terme *chance* ne semble pas poser de problème. De cette manière, le contexte aiderait à la compréhension d'idées sans nécessité de faire recours à des termes précis, nous reviendrons plus tard sur ce sujet. De plus, étant les aspects calculatoires corrects, l'enseignant ne rejette pas son utilisation.

L'approche commence donc par l'énonciation des cas favorables et possibles, en d'autres termes à partir de la dimension numérique. La valeur de la probabilité reste pour les élèves sous la forme d'une proportion[15], l'enseignant, lui, a tendance à utiliser la formulation décimale.

Une fois énoncés les éléments principaux de l'algorithme de calcul (cas favorables et possibles), un élève établit la relation entre les deux dimensions de la probabilité. Le graphique, étant toujours projeté sur le mur principal de la salle, il n'est pas considéré dans son analyse. Dans sa mise en rapport des deux dimensions, il s'appui sur l'échantillon des deux cent vingt lancers pour indirectement introduire la notion de probabilité comme une moyenne :

Donc on a deux cent vingt tirages ça fait cinquante cinq
Alors çà veut dire quoi ton cinquante cinq tu peux l'interpréter
Le nombre de possibilités qu'on va avoir de trouver deux fois d'avoir un sur le nombre de tirages fait
(…)
Ouais à peu près
A peu près

Son point de départ est la proportion calculée (une chance sur cinq), en suite il la projette sur l'échantillon de deux cent vingt cinq lancers et détermine le nombre de cas favorables attendu pour cet échantillon. Le sens du mot "attendu" étant implicite, il en fait référence de manière indirecte : « **Ouais à peu près** ».

[15] Certains auteurs proposent pour la probabilité le format de pourcentage comme le plus adapté aux représentations des élèves ((Gigerenzer, 1994)).

Cet élève utilise donc non pas la notion d'attente à long terme ni de fréquence d'apparition mais plutôt celle d'attente moyenne. Au lieu de considérer la série infinie et sa convergence (probablement trop abstraite pour lui), il se sert d'une représentation plus facilement saisissable, la proportion attendue. Cette proportion se contextualise et devient le cardinal attendu (cinquante cinq) pour l'échantillon en question (deux cent vingt cinq).

Il semblerait plus accessible et probablement plus significative aux élèves la représentation de *cardinal attendu* par rapport à un total de référence au lieu de la stabilisation de fréquences. La modalité de choisir des cardinaux a été déjà utilisée à la minute trente cinq lorsque est apparu le terme *moyenne*. A l'occasion, un élève l'a utilisé pour se référer au nombre de coups attendus pour gagner la partie (*Chapitre III. De l'évaluation comparative vers la numérique*).

D'ailleurs, le graphique projeté n'a pas surpris à personne. Il semblerait que la notion de stabilisation de fréquences dans ce genre de contexte (jeu de hasard) serait déjà incorporée chez les élèves à cet âge. Cette notion pourrait se trouver instable, nous l'ignorons. Néanmoins, et aux fins du problème, il parait une notion déjà acquise. Cette impression a été partagée par l'enseignant qui décide de continuer avec le déroulement[16], même si nous avions prévu dans notre planning une répétition de l'expérimentation afin d'augmenter la taille de l'échantillon pour une meilleure étude de la convergence.

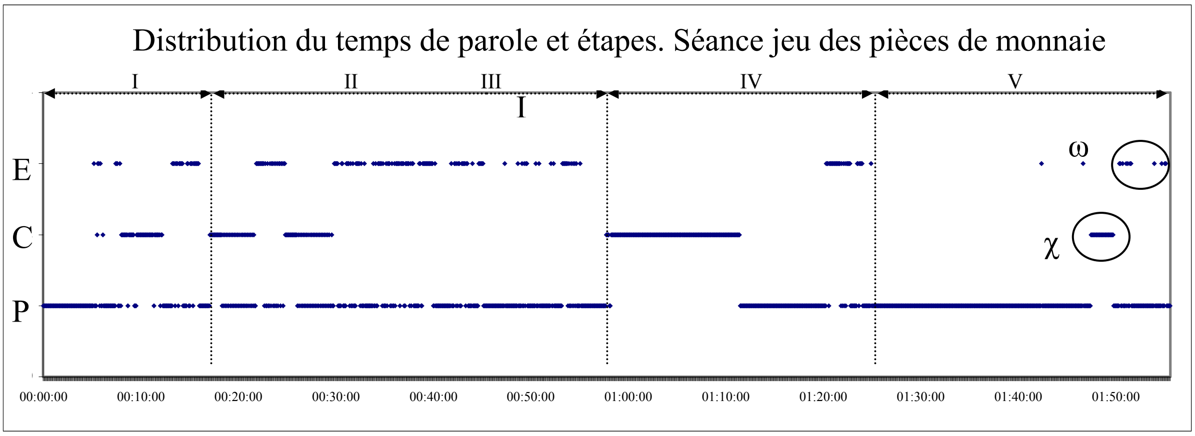
Après quelques échanges autour du cardinal attendu dans l'échantillon de deux cent vingt quatre lancers (minute quatre vingt quatre) la parole devient presque exclusive de l'enseignant. A partir de ce moment, toute notion relative à l'étude des tendances appartient à l'étape d'institutionnalisation. Cela ne veut pas dire que le débat s'est épuisé entre élèves, mais plutôt une conséquence du virement de dynamique de la séance. A partir de ce moment donc, l'enseignant commence une nouvelle étape qui nous avons dénommée institutionnalisation, principalement par la place que l'enseignant accorde au bilan des notions traitées dans cette dernière étape.

Institutionnalisation

Cette étape est, naturellement, de forte présence discursive de l'enseignant, le contraste entre les étapes précédentes et celle-ci est évidente. Le Graphique 41 montre la

[16] Lors de la séance, l'enseignant s'est adressé à nous pour nous communiquer qu'il ne serait pas nécessaire de répéter l'expérimentation, vue l'apparente compréhension des élèves.

distribution de parole de cette dernière partie de la séance. Dans ce graphique nous identifions trois instances, la première dont l'enseignant s'adresse aux élèves pendant plus de vingt trois minutes, la deuxième (L) dont les élèves effectuent des calculs de probabilité et la dernière (M) dont se produisent des échanges autour du critère décisionnel, nous commencerons par l'analyse de la première partie de cette étape d'institutionnalisation.



Graphique 41

La place des échanges des élèves

Dans la première partie de cette dernière étape (jusqu'à l'instance χ) l'enseignant procède à institutionnaliser un certain nombre de concepts. La modalité implémentée dans cette phase de mise en commun pourrait se décrire en général comme étant composée de cycles, dont un pour chaque concept. L'enseignant commence par l'énonciation du concept de manière relativement formelle et puis il présente une illustration à partir du jeu de la séance. De cette manière sont présentés des concepts tels que hasard, expérience aléatoire, expérience équiprobable, cas possibles et favorables, univers, loi de probabilité, formule de Laplace et gain moyen. Par exemple, le cycle correspondant à la notion d'univers est :

90 *(…) donc l'univers qu'on appelle oméga l'univers c'est l'ensemble oméga*
91 *De tous les résultats possibles par exemple c'était quoi notre univers c'était tous les résultats possibles qu'on peut obtenir si on lance deux pièces première pièce pile deuxième pièce face deuxième pièce pile pardon première pièce pile deuxième pièce pile on peut obtenir face face aussi on peut obtenir face pile donc nos résultats possibles ici dans notre exemple çà serait oméga pile pile pile face face pile face face (…)*

Ces cycles sont bipolaires, un pôle est la formulation du concept, l'autre l'illustration sur la situation. En d'autres termes, l'enseignant pour cette étape d'institutionnalisation ne reprend pas dans son discours les précédents échanges des élèves. La structure de ces cycles met en évidence :

- l'importance des caractéristiques contextuelles du problème pour le traitement d'un concept. En particulier pour la probabilité, et pour que l'enseignant puisse s'en appuyer, le contexte doit se correspondre avec la notion visée.
- des difficultés à institutionnaliser une notion (au moins ces notions) sur la base des échanges des élèves. L'absence d'échanges des élèves dans l'étape d'institutionnalisation dans cette séance en particulier pourrait s'expliquer en partie par un effet de nouveauté. Cette séance fut, pour l'enseignant, sa première expérimentation avec nous et sur cette situation en particulière. A fin de mieux gérer cette nouvelle situation il s'est préparé à l'avance une série de notes contenant les notions à institutionnaliser. Ces notes projetées sur le mur, étaient d'utilité double, d'une part elles lui servaient d'appui pour la mise en commun, d'autre part, elles pouvaient être recopiées par les élèves sur leurs cahiers. Il est évident que dans ces notes il manquait toutes les opinions versées par les élèves lors des débats. Probablement dans un cadre de classe moins inhabituel, l'enseignant inclurait ces opinions lors de l'institutionnalisation.

Pour ce qui concerne l'institutionnalisation de la probabilité, il y a eu un écart entre le discours des élèves et celui de l'enseignant dans cette étape, et cela tant dans les dimensions sémantique que calculatoire. L'interprétation est institutionnalisée en tant que tendance à long terme, notion directement liée au graphique de référence mais quelque part éloignée de celle des élèves qui se sont exprimé en termes de attente moyenne. La représentation numérique a été elle aussi différente, l'enseignant formalise la représentation décimale de la probabilité tandis que les élèves ont fait référence à des proportions.

D'ailleurs, certains choix de l'enseignant semblent s'expliquer les uns par ses habitudes en tant qu'enseignant, les autres par ses responsabilités pédagogiques et institutionnelles, toutes les deux interdépendantes. Bien que l'enseignant soit resté toujours attentif aux besoins de notre recherche, il a pris certaines décisions en fonctions de ces deux axes. Pour ce qui concerne ses habitudes (dont nous y incluons ses représentations), cet axe expliquerait la manière dont il résout l'institutionnalisation de la espérance mathématique, notion certes au cœur du critère décisionnel du problème. Pour ce qui concerne ses responsabilités pédagogiques et institutionnelles, elles expliqueraient le genre de notions institutionnalisées dans la dernière étape de la séance.

Quelques habitudes de l'enseignant

La place de l'espérance mathématique dans l'étape d'institutionnalisation a été l'objet de réflexions lors de nos entretiens. Ces échanges se sont produits de manière imprévue et malheureusement nous ne disposions pas à l'occasion des outils pour son enregistrement. Toutefois, ce que nous avons retenu du sujet semble correspondre avec l'observé lors de la séance. L'enseignant, à l'entretien, nous avait fait savoir que la notion d'espérance mathématique ne faisait pas partie des objets habituellement traités au début du chapitre Probabilités, et cela principalement parce que sa présentation requert de la notion de variable aléatoire. Cette dernière lui semblait trop prématurée de l'introduire en la première séance. Il nous avait proposé donc de ne pas formaliser l'espérance mathématique et d'en retenir seulement les aspects sémantiques. Cette décision explique que le terme retenu ait été *gain moyen*, et que son expression mathématique n'ait pas été que verbalement évoquée, sans aucune représentation symbolique sur le tableau.

Parmi ses habitudes (y compris représentations) se trouverait donc le requit de, lorsqu'une notion est institutionnalisée, elle doit se faire de manière à que tous ses composants soient susceptibles d'être institutionnalisés aussi. Autrement dit, il ne serait pas possible ou au moins convenable d'institutionnaliser partiellement un concept. Cette habitude expliquerait la présentation donnée à l'espérance mathématique par une approche plutôt sémantique, en contournant pour le moment sa formulation symbolique et sa dénomination.

Des responsabilités institutionnelles

Les responsabilités institutionnelles qui pèsent sur l'enseignant pourraient expliquer d'autres actions entreprises par l'enseignant. Il était de notre intérêt faire valoir la probabilité comme un outil de prise de décision, une probabilité devenant un élément d'une rationalité pour agit face à l'incertain, cela explique l'importance que nous attribuons à l'interprétation de la probabilité. En autres termes, il n'est pas si nécessaire de s'intéresser aux interprétations si on focalise le travail sur des aspects calculatoires, pratiquement rien ne change en termes calculatoires pour l'une ou l'autre des interprétations. C'est lorsqu'on considère la place de la probabilité dans des critères rationnels de prise de décision que son interprétation devient importante.

S'intéresser en classe au rôle de la probabilité dans une rationalité de l'incertain conduit directement à s'intéresser à l'indéterminisme, et comment en agir. Ces aspects semblent avoir été relégués ou au moins régulés par rapport aux notions habituellement enseignées en BTS. Des notions concernant des prises de décisions telle que le critère de maximisation de l'espérance ne font pas partie des notions habituellement enseignées, (nous a confié l'enseignant). Ainsi, ces concepts, étant plus d'intérêt pour notre recherche que pour le programme de l'enseignant, ils n'arrivent pas au même degré de formalisation que ceux présents dans le programme.

Dans cette dernière étape d'institutionnalisation il y a deux moments où la parole est rendue aux élèves (χ et ω, Graphique 41). Le premier de ces moments, des échanges au sein des groupes, se produit à partir de la minute cent sept. Ce changement dans la gestion de la parole correspond à la demande de l'enseignant de déterminer le gain moyen pour le choix de trois pièces de monnaie. Vu que pour deux pièces de monnaie le lien entre les dimensions calculatoire et la sémantique de la probabilité n'a pas posé des problèmes, l'enseignant fait que la détermination de ce gain moyen s'effectue entièrement sur le plan calculatoire, en d'autres termes sans faire recours à une expérimentation. Le moment ω dans le graphique correspond donc aux échanges d'opinion entre élèves tendant à déterminer non seulement la probabilité de gagner des points en jouant avec trois pièces mais aussi son gain moyen. Cette tâche se réalise en moins de trois minutes. Vu que la fin de la séance approche, l'enseignant accélère la cadence et rappelle la technique de calcul du gain moyen pour deux pièces, ils l'appliquent, mais les résultats pour le dénombrement de cas possibles sont différents entre eux («**huit** », « **six** », « **dix** », ω, Graphique 41).

L'enseignant, ayant institutionnalisé les concepts mais non pas les techniques de calcul, se voit forcé à intervenir. Il revient sur la méthode de calcul et décline son explication en précisant toutes les combinaisons possibles pour les trois pièces. De cette manière, de la main de l'enseignant, la classe arrive à la valeur de la probabilité d'obtenir trois fois pile. Un élève anticipe en suite, en estimant, que le gain moyen sera identique aux précédents: «**Toujours la même chose** ». Cette nouvelle proposition est vitement écartée, le gain moyen leur donne trois virgule soixante quinze ; ce calcul est effectué toujours sous les directives de l'enseignant.

Sur les interprétations et la prise de décision

A moins de deux minutes de la fin de la séance et en le sachant, l'enseignant demande d'interpréter la valeur du gain moyen. Après une erreur d'un élève (« [C'est] **La probabilité** ») un autre corrige (« **La moyenne** »), finalement l'enseignant reprend à sa charge l'interprétation de la valeur trouvée, dans son explication il se sert du signifié institutionnalisé, la tendance à long terme. En suite, pour quatre pièces, les calculs s'accélèrent :

Un sur seize
115 *Un sur seize donc quarante*
Quarante sur quarante fois un sixième
Quarante sur seize et çà fait deux virgule cinq bon alors vous vous prononcez comment alors vous vous prononcez comment alors qu'est-ce qu'il faut prendre
Une pièce
Une pièce
Une pièce
Oui c'est cinq

Finalement, les élèves entendent le signal de la fin de la séance, l'enseignant de sa part, essaie que les élèves répondent à la question du problème ; qu'elle est la stratégie la meilleure :

Attendez attends attends attends je n'entend rien là
Une pièce parce que c'est cinq le gain
Quand il y a deux pièces c'est combien le gain
Eh cinq
C'est cinq aussi
une ou deux pièces
Eh oui on peut prendre une ou deux pièces par contre si on en prend quatre on va moins gagner (...)

La dynamique de ces dernières minutes ne peut s'inscrire que dans la nécessité de l'enseignant de conclure le problème avant la fin de la séance. Ses prises en charge et ses indices glissés, tentent de accélérer la cadence de la séance pour arriver à faire prononcer les élèves avant que le signal sonne.

Notons que même si il est pressé, l'enseignant souhaite que les résultats des calculs soient replacés dans le contexte du problème, autrement dit, il cherche à que le problème soit conclu dans ces aspects sémantiques et non seulement calculatoires, mais la classe est trop agitée pour s'y concentrer et la contextualisation des calculs reste finalement vague. Dans cette séance il n'y a pas eu le temps de, une fois les calculs effectués, débattre sur la stratégie retenue, et de notre part, d'observer le degré d'acceptabilité d'une telle proposition. Une heure plutôt (minute quarante quatre) et lorsque la stratégie de la moyenne attendue commençait à s'esquisser, un élève s'est opposé à décider sur ses lancers à partir d'une "notions théorique" telle que le gain moyen : « **Oui mais là tu peux pas te fier sur laquelle tu ne peux pas te fier pour faire tes lancers** ». Il semblerait qu'entreprendre une action fondée sur un critère décisionnel relève des questionnements plus profond qu'un calcul mathématique. Entreprendre une action de ce genre et la soutenir par des arguments du type traité lors de la séance impliqueraient de profonds changements dans les conceptions déterministes des élèves et sur les critères de validation couramment utilisés. Malheureusement nous n'avons pas eu la possibilité d'observer de tels débats à la fin de la séance.

Néanmoins, et malgré l'empressement, l'enseignant réserve quelques instants à l'interprétation du gain moyen calculé (et pour autant à la probabilité concernée) avant de centrer l'attention sur la stratégie gagnante. L'ordre de cette séquence, interprétation de la probabilité, puis prise de décision illustre ce que nous avons dénommé le caractère nécessaire de l'interprétation de la probabilité pour une prise de décision (*Chapitre I* et *Chapitre II*). Même si l'explication se précipite à cause de la fin de la séance, elle se déroule dans les deux temps indiqués, et nous ne nous référons pas ici à une chronologie dans cette séquence (d'abord interprétation puis prise de décision), mais plutôt à l'indéfectible évocation de l'interprétation de la probabilité lorsque l'on développe les arguments d'une décision. Par la suite nous présenterons les conclusions pour cette première expérimentation.

Conclusions

Sur les éléments caractéristiques

La nature de l'objet

Pour cette situation-problème les objets sur lesquels portent les probabilités ne sont pas toujours de nature fréquentiste. Nous avons identifié trois objets susceptibles d'être probabilisés :

- *Séries infinies*. Elles correspondent à une succession de tirages. Par exemple,

  P(A) où l'événement A est obtenir pile lorsqu'on lance (un grand nombre de fois) une pièce.

  P(A) signifie la fréquence d'apparition de l'événement « obtenir pile » lorsqu'on répète l'expérience jusqu'à l'infini.

  Cette probabilité a été sémantiquement évoquée et quantifiée lors de la séance (min : 35 « **c'est une moyenne** », min 42: « **une infinité de tirages**», min 42 : «**beaucoup plus de tirages** », etc.)

- *Epreuves génériques*. Elles correspondent à une proposition susceptible de reproduction. Par exemple,

  a) P(B) où l'événement B est la probabilité d'obtenir cinq fois (pile, pile) en dix lancers de deux pièces.

  P(A) signifie le degré de certitude d'obtenir cinq fois (pile, pile) dans une série de dix lancers.

  Cette probabilité a été sémantiquement évoquée (min 37 : «**Mais sur dix lancers on peut pas savoir** ») mais non numériquement évaluée.

  b) P(C) où C est la probabilité de réussir la partie (en gagnant les cinquante points).

  P(C) est le degré de certitude de gagner la partie en jouant contre un adversaire.

  Cette probabilité a été sémantiquement évoquée (min 14 : « **plus il y a de pièces plus la chance de réussir est faible** »), son évaluation numérique par des ensembles de référence finis est impossible, elle reste qualitative.

Les probabilités bayésiennes évoquées lors de la séance (épreuves génériques) n'ont pas été calculées, soit par son impossibilité (item (c) soit par sa non relevance dans le contexte de son évocation (item (b). La seule a avoir réunie les deux dimensions (sémantique et numérique) a été la fréquentiste, notion visée pour la séance.

D'ailleurs, l'enseignant était au courant de nos intentions de sensibiliser, pour cette séance, la notion fréquentiste. Dans les entretiens ayant précédé la séance nous avons choisi, en cas de manifestation d'une probabilité bayésienne, de ne pas encourager son calcul, restant ainsi de caractère qualitatif.

Type de raisonnement et valeurs logiques

Pour aboutir à déterminer la valeur numérique des probabilités fréquentistes du jeu il y a couramment deux moyens, soit on admet des hypothèses d'équipossibilité des événements élémentaires, soit on l'estime par un échantillon suffisamment important.

Etant donné nos suppositions de notions élémentaires du calcul de probabilité déjà acquises au lycée et vu nos objectifs de relier ces deux moyens d'évaluation, nous avons choisi pour cette séance d'articuler les deux approches de la manière suivante : le calcul (*a priori*) s'effectuerait, sous le prétexte d'économie de temps et puis la stabilisation de fréquences aurait lieu en tant que vérification.

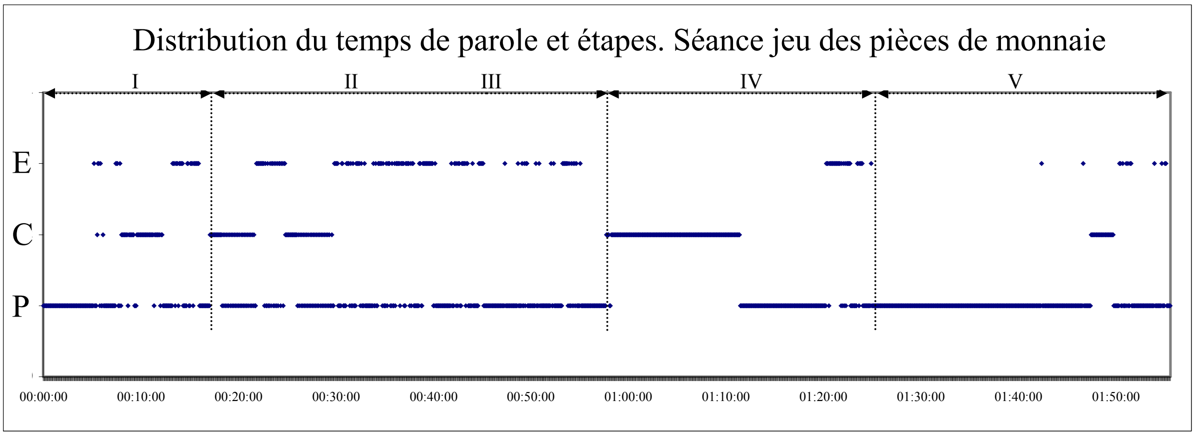
Pour le calcul établi *a priori* le type de raisonnement mobilisé est le déductif, à partir de l'admission de l'équipossibilité se déduisent les valeurs d'attente à long terme pour chacun des événements concernés. Ce type de raisonnement suffit pour calculer une probabilité, mais non pas pour établir un critère décisionnel.

Lors de la séance, nous avons observé le phénomène de repli du groupe dénommé "indéterministe". Ce groupe, qui s'est intéressé dans un moment donné (minute 42 de la séance) à la moyenne de points gagnés, revient sur ses pas, à cause des critiques du groupe "déterministe" et se limite enfin à proposer des tâches calculatoires. Ces tâches correspondent aux étapes du calcul de la probabilité *a priori* de gagner avec deux pièces. C'est ce type de raisonnement (déductif) qui fini par dominer le débat, bloquant la construction d'un critère décisionnel.

Un critère tel que la maximisation de la espérance mathématique ne garantit pas la réussite ; en d'autres termes il n'assure pas la valeur logique vraie pour la proposition "je gagnerai la partie". Ainsi, la classe n'avance pas sur sa construction car il ne s'ajuste pas au paradigme dominant la séance, c'est l'enseignant qui doit intervenir donc pour le valider comme critère décisionnel. Le conflit entre paradigmes a été donc une de plus importantes difficultés dans la résolution de ce problème.

Sur l'interprétation de la probabilité

Dans cette séance, l'interprétation de la probabilité en tant qu'attente à long terme a été entendue presque exclusivement de la part de l'enseignant, qui lors de l'institutionnalisation (Etape V. Graphique 42), présente une synthèse des notions centrales de la séance. Avant et après ce moment, il l'évoque par des termes assez proches de ceux des élèves (chance, possibilités) ; cela s'expliquerait par son intention de s'approcher au discours des élèves et ainsi faciliter l'échange d'idées.



Graphique 42

Les élèves de leur part, et pour ce qui concerne la probabilité, sauf une exception (« **si si sur un nombre infini de lancers** ») ils ont recours principalement aux termes *chance* et *moyenne,* chacun retenant des aspects différents de la probabilité. Le premier (chance) est utilisé plutôt pour évaluer la probabilité, généralement l'ensemble de référence (fini) est aussi évoqué (« **une chance sur quatre** »), sousjacent à cette utilisation des hypothèses d'équipossibilité de cas élémentaires. Il est usuel aussi que ce soit accompagné d'un quantificateur (« **plus de chances** »), cette formulation (chance + quantificateur) est typique de l'étape d'évaluation comparative dont les événements sont confrontés par paires, la comparaison s'établit par l'usage des deux types d'ensembles de référence repérés dans le Chapitre I : fini (« **une chance sur quatre** ») et infini (« **plus rare** »).

Le terme *moyenne*, à la différence de *chance*, est plus proche des aspects sémantiques de la probabilité. Lors de la séance il est utilisé pour souligner la compensation qui se produit entre les résultats des tirages successifs des pièces de monnaie. L'existence d'une (unique) valeur pour cette moyenne semble être acceptée pour les élèves, au moins pour ce genre de contextes. Les échanges se basent sur la supposition de que cette moyenne est unique. De cette manière, l'admission de l'existence d'une valeur pour la moyenne indiquerait que les élèves acceptent sa stabilisation de fréquences.

Conflits entre paradigmes

Les débats suggèrent donc que les principes de la notion de probabilité fréquentiste pour ce genre de contextes (ludiques, d'équipossibilité de cas élémentaires) seraient déjà acquises chez ces élèves, tant dans ses aspects calculatoires que sémantiques. Néanmoins, même la notion de probabilité déjà acquise, les élèves, sans le guide de l'enseignant, ne sont pas arrivé à se prononcer sur la stratégie gagnante. Les raisons seraient multiples, nous souhaitons retenir en particulier celle des conflits de paradigmes.

En effet, lors des débats, les élèves ne manifestent pas de doute sur les notions élémentaires autour de la probabilité (stabilisation de fréquences, equipossibilité de cas élémentaires, correspondance entre calcul *a priori* et vérification par stabilisation de fréquences, etc..). Leurs désaccords concernent la stratégie gagnante et les arguments la validant, il semblerait que personne ne doute que lorsqu'on lance une pièce la fréquence d'apparition de la face "pile" est de cinquante pour cent des tirages « **en moyenne** ». Ce qui est objet de discussion est en tout cas, en quoi cela pourrait contribuer à réussir une partie.

Lorsque un groupe d'élèves, en se servant de la notion de probabilité, propose de s'intéresser au gain moyen, un autre conteste cette approche en argumentant l'impossibilité d'assurer la réussite avec cette nouvelle approche. Le motif des discussions ne serait pas donc la notion de probabilité en soit mais plutôt un conflit d'approches pour agir devant l'incertain, en d'autres termes, un conflit entre deux paradigmes, l'un caractérisé par des valeurs logiques dichotomiques, l'autre par des valeurs intermédiaires de certitude.

Le premier valide ses propositions par déductions, le deuxième par des inférences diverses telles que l'induction ou l'abduction. Ces deux positions nous les avons dénommées comme étant l'une "déterministe" et l'autre "indéterministe". Les élèves placés dans la première approche semblent n'agir qu'en cas de certitude, ils n'admettent pas comme argument qu'une stratégie soit plus convenable qu'autre parce qu'elle apporte plus de "chances" de réussir le jeu. Les élèves "indéterministes", par contre construisent leurs propositions sur les "chances" de réussir, même si la certitude de réussite n'est pas assurée.

Dans cette séance les élèves "indéterministes" ont cédé face aux critiques des "déterministes". Les premiers, s'etaientr repliés jusqu'à même se contourner à des propositions déterministes, finissant par abandonner l'intention de soutenir leurs arguments. L'enseignant se voit donc forcé d'intervenir pour valider l'approche et ainsi rouvrir le débat vers la solution de la stratégie de l'espérance maximale.

A l'origine de la résistance à quitter le paradigme déterministe se trouveraient au moins deux facteurs, le premier d'ordre épistémologique, des difficultés propres à se positionner dans un nouveau paradigme dont les critères de validation sont bien différents. Le deuxième d'ordre didactique, en classe de mathématiques le contrat didactique, construit au but de plusieurs années, tend à favoriser les raisonnements déductifs plutôt que les inductifs.

Pour ce qui concerne les difficultés épistémologiques, plusieurs chercheurs ((Dale, 1999; Hacking, 2002), (Shafer, 1996)) ont signalé que lors de l'émergence de la probabilité, le grand saut conceptuel n'était pas autant la notion de probabilité en soi[17] mais plutôt en la rationalité d'un système pour aborder l'incertain. Il est vrai que la probabilité en est une pierre angulaire, mais il semble que la difficulté épistémologique majeure fut d'en construire une nouvelle approche, dont les propositions ne se valident pas par leur certitude mais par la rationalité des arguments.

Il semble se reproduire cette difficulté épistémologique dans la séance : les élèves, connaissant déjà la notion de probabilité, eprouvent des difficultés à construire un ensemble de critères pour décider rationnellement devant l'incertain.

Pour ce qui concerne les obstacles didactiques, bien que nous n'avons pas testé cette expérimentation dans une classe autre que de mathématiques, et que nous ignorons pour autant les effets qu'un tel changement de cadre disciplinaire provoquerait chez les élèves, il semble néanmoins que des aspects didactiques pourraient avoir contribué à décourager les élèves "indéterministes" à poursuivre leur stratégie.

En classe de mathématiques le paradigme le plus favorisé est le déterministe, les deux groupes d'élèves se trouveraient donc en situation inégale pour soutenir leurs arguments. Les uns savent qu'il est impossible d'assurer la réussite, et que toute stratégie peut être "réfutée" par un contre exemple ; ils sont donc outillés pour agir. Leur position correspond avec l'approche le plus favorisée en classe de mathématiques, le déterminisme. Pour eux, aucune stratégie n'étant certaine, il n'en existerait pas une gagnante.

Pour les autres, qui soutiennent l'existence de stratégies plus convenables que d'autres, leur regard est tout autre, leur choix se fonde sur ce qu'il arrive lorsque l'on répète l'expérience, la réussite n'étant pas assurée, l'argument se défend par sa cohérence. A la différence de l'approche déterministe qui assure la validité de la proposition, celle-ci ne peut assurer que sa rationalité, elle doit donc être soumise à une consensus pour son approbation ; ce consensus devient difficile à se consolider pour la dominance du déterminisme en classe de

[17] Dont des cultures connaissaient déjà comment retoucher un dé pour le rendre de faces équiprobables ((Hacking, 2002))

mathématiques, c'est l'enseignant qui finalement prend en charge la validité de cette rationalité.

D'autres conclusions seront proposées à la fin du chapitre, lors de la présentation des deux autres expérimentations réalisées avec le même groupe d'élèves. Nous analyserons la deuxième situation-problème.

4.1 Analyse de deuxième expérimentation

Nous avons identifié trois types d'objets sur lesquels peut porter une probabilité (*Chapitre I. Une première différence : Les objets* A en P(A)). Un de ces objets correspond à la notion fréquentiste de la probabilité, nous l'avons dénommé série infinie ou *collective* (von Mises). Les deux autres correspondent à la notion bayésienne, dont un est l'épreuve générique et l'autre une hypothèse. Le problème des pièces de monnaie correspondait à la notion fréquentiste, ce deuxième s'adresse à sensibiliser sur la notion bayésienne de la probabilité en sa modalité épreuve générique. Nous commençons cette analyse par la présentation de dette situation-problème.

Les circuits

Analyse a priori

Matériel :

Sept enveloppes opaques. Trois contenant des fusibles ( ), deux des résistances ( ) et les deux derniers des interrupteurs ( ) (Figure 20). Dans toutes les enveloppes il y a un certain nombre de pièces défectuesses (voir Tableau 57). Finalement, une pile et une ampoule.

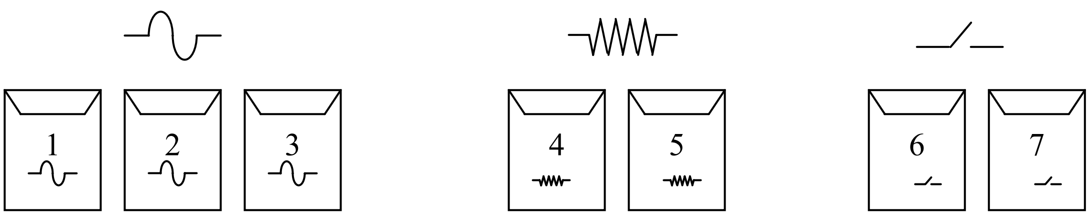


Figure 20

N° pièces	Env. 1	Env. 2	Env. 3	Env. 4	Env. 5	Env. 6	Env. 7
Fonctionnant	6	2	7	7	8	2	7
Défectuesses	1	5	1	2	1	6	1
Total	7	7	8	9	9	8	8

Tableau 57

Le problème :

En se servant des pièces à l'intérieur des enveloppes, l'enseignant construit trois circuits ouverts. Pour cela, des enveloppes 1 et 4 il prend au hasard un fusible de la première et une résistance de la deuxième, en les connectant il obtient l'assemblage A. De la même manière il construit l'assemblage B à partir des pièces des enveloppes 2 et 6, finalement le C en prenant au hasard les pièces des enveloppes 3, 5 et 7 (Figure 21).

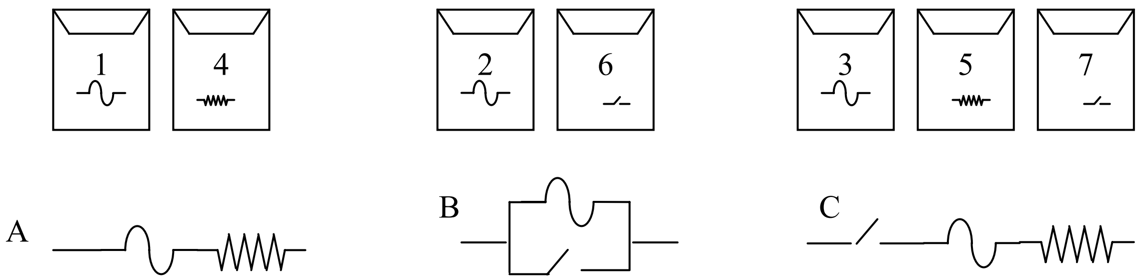


Figure 21

Il montre à la classe le circuit ouvert de la Figure 22 composé de l'ampoule et la pile, et pose la question suivante : je souhaite allumer l'ampoule, quel de ces trois assemblages (A, B ou C) me conseilleriez-vous d'utiliser pour fermer le circuit ?

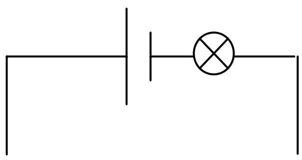


Figure 22

Objectif pour l'expérimentation

Avec cette expérimentation nous nous intéressons aux conditions et possibilités de traitement de la notion bayésienne de la probabilité, en particulier au genre dénommé épreuve générique. Ces épreuves, rappelons-le, se caractérisent par leur unicité et en même temps par la possibilité de reproduction.

Dans ce problème la question que pose l'enseignant se réfère au fonctionnement des assemblages A, B et C, présentes dans la salle devant les élèves. Ces assemblages, même si uniques, deviennent génériques par le manque d'information particulière les concernant.

Les épreuves génériques ce sont les objets les plus délicats à saisir, leur caractère générique et leur possibilité de reproduction conduisent immédiatement à penser à la fréquence d'apparition. De cette manière, un artifice sur le plan rhétorique peut conduire à une substitution d'objet, remplaçant l'épreuve générique par la série infinie. Avec la présence explicite des objets (les assemblages) nous souhaitons contribuer à fixer l'objet sur lequel porte la probabilité. En d'autres termes, si les élèves, dans la résolution du problème, évoquent la fréquence d'apparition, l'enseignant pourra s'appuyer sur les assemblages présents dans la salle pour bien préciser que la question concerne les circuits placés devant eux et non pas ce qui arriverait lors d'éventuelles répétitions de l'expérimentation. La présence matérielle est donc une sorte de variable didactique ayant pour but de faciliter la fixation de l'objet épreuve générique.

Pour cette expérimentation nous nous intéressons à observer si une telle notion de la probabilité est possible d'être cerné en classe tant dans ses aspects sémantiques que calculatoires. Pour l'occasion et afin de minimiser ce risque, nous avons choisi pour son évaluation numérique, un ensemble de référence fini au lieu de l'infini qui ferait intervenir explicitement la fréquence d'apparition.

D'ailleurs, et de la même manière qu'en la première expérimentation, la résolution du problème des circuits requiert une articulation entre le modèle mathématique et l'interprétation de la situation réelle. En fait nous considérons, dans ce genre de situations, la modélisation comme étant juste une partie du problème, l'autre concerne l'interprétation. Cette interprétation inclut non seulement le contexte de la situation mais aussi celle des objets du modèle. La Figure 23 représente ce schéma, dans lequel résolution d'un problème implique d'abord un travail interprétatif et puis un de modélisation. L'intention de cette figure est de souligner que l'interprétation précède l'acte de modélisation, l'accompagne, le guide et le suit. En effet, l'interprétation du problème doit précéder l'acte de modélisation, autrement, la convocation d'objets tels que la probabilité ne serait possible que comme conséquence du contrat didactique ou par indication explicite de l'enseignant. L'interprétation doit aussi accompagner la modélisation, si le lien entre les deux n'est pas établi, les objets mathématiques n'ont pas de raison d'être. Finalement, l'interprétation doit suivre la modélisation pour réintroduire dans le contexte les transformations effectuées dans la modélisation et ainsi, dans ces cas, pouvoir argumenter les décisions.

Un des moments les plus délicats dans ce processus constitue l'étape de modélisation. Premièrement parce qu'en classe de mathématiques il n'est pas tout à fait habituel d'habiller un problème par un contexte réel, l'analyse des manuels nous l'a montré. Deuxièmement, parce que cette étape implique un double effort, d'une part travailler sur le plan des objets symboliques (par exemple), muni de règles de transformation et de représentation, et d'autre part accompagner sémantiquement les objets manipulés et leurs transformations respectives.

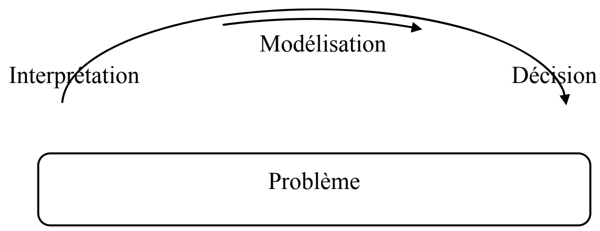


Figure 23

Pour ce qui concerne la probabilité en particulier, le moment délicat est celui des opérations mathématiques dont une probabilité est numériquement transformée par les règles du registre symbolique. Cette transformation sur le plan abstrait mathématique peut entraîner une erreur consistant à interpréter une probabilité par exemple d'une manière fréquentiste et puis, après une transformation numérique, lui assigner l'interprétation bayésienne. Cette erreur nous l'avons décrit dans le *Chapitre I. Définitions sémantique et opératoire,* que nous reproduisons ci-dessous.

Erreur par échange d'interprétation postérieur aux calculs

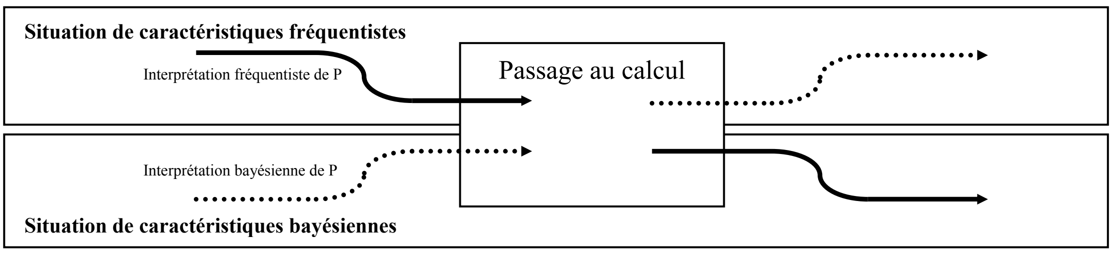


Figure 6

Dans ce graphique, une situation-problème mènerait à se représenter, par ses caractéristiques contextuelles, une des approches de la probabilité et puis, postérieurement aux calculs, son signifié serait échangé, entrant ainsi en contradiction avec les caractéristiques du problème.

D'ailleurs, la deuxième situation-problème que nous avons proposée à l'enseignant fait intervenir plusieurs probabilités, d'une part les probabilités de fonctionnement des composants et d'autre part celles des assemblages, des objets tous différents mais de même nature.

Selon les méthodes utilisées par l'évaluation numérique, la démarche du calcul des probabilités des assemblages sera plus ou moins complexe. Dans un extrême de complexité se trouve celle qui détermine la probabilité des assemblages à partir des probabilités de leurs composants. Dans l'autre, celle qui le fait par un simple dénombrement d'assemblages possibles et favorables.

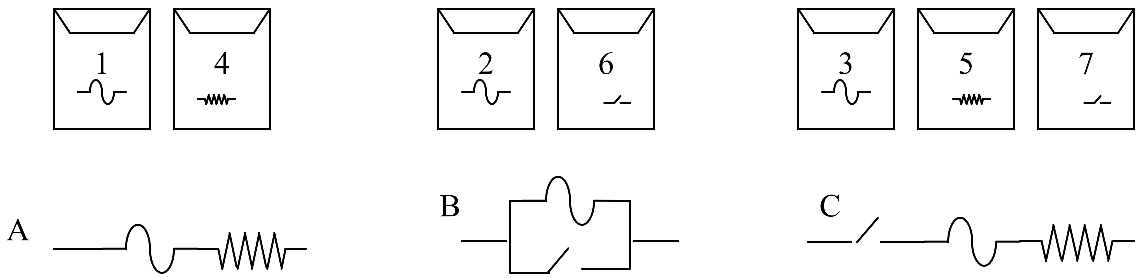
La première de ces démarches est complexe. Deux transformations se produisent en même temps, l'une numérique et l'autre d'objet. En effet, en partant des probabilités des composants on arrive à celle de l'assemblage, il y a donc une transformation sur le plan numérique en passant d'une valeur à une autre. Mais il y a aussi une transformation d'objet, lorsque l'on commence le calcul par les composants et termine par l'assemblage. Néanmoins, ces deux transformations n'impliquent pas une d'interprétation, toutes ces probabilités signifient des degrés de certitude du fonctionnement de leurs objets respectifs. Un autre motif de difficulté consiste en l'opération mobilisée pour la transformation numérique. Cette opération, le produit, doit s'expliquer sémantiquement en relation au problème pour qu'elle puisse se justifier, tâche qui pourrait s'avérer difficile aux élèves.

La deuxième de ces démarches, celle qui détermine les probabilités de fonctionnement des assemblages par le dénombrement d'assemblages possibles renvoie aux principes mêmes du calcul de la probabilité, elle est donc la plus simple à saisir en termes tant sémantiques que calculatoires. Nous analyserons ces possibles solutions dans le déroulement ci-dessus détaillé.

Solution :

Dans un point de vue bayésien, le problème se résout en choisissant l'assemblage pour lequel sa probabilité de fonctionnement est la plus élevée. Le calcul des probabilités du fonctionnement des assemblages se fait en détaillant tous les assemblages possibles soit à partir des probabilités de leurs composants. Dans les deux cas les évaluations se réalisent par référenciation (*Chapitre I. La référenciation*) d'ensembles finis.

Nous présentons par la suite la solution numérique du problème. Les aspects sémantiques et les moyens possibles mobilisés par les élèves seront développés au fur et à mesure que nous déployons le déroulement possible.



N° pièces	Env. 1	Env. 2	Env. 3	Env. 4	Env. 5	Env. 6	Env. 7
Fonctionnant	6	2	7	7	8	2	7
Abîmées	1	5	1	2	1	6	1
Total	7	7	8	9	9	8	8

Soient les propositions :

A : « l'assemblage A fonctionne », B : « l'assemblage B fonctionne » et C : « l'assemblage C fonctionne » :

$$P(A) = \frac{6}{7} \times \frac{7}{9} = \frac{2}{3} = 0{,}67$$

$$P(B) = \frac{2}{7} + \frac{1}{4} - \frac{2}{7} \times \frac{1}{4} = \frac{13}{28} = 0{,}46$$

$$P(C) = \frac{7}{8} \times \frac{7}{8} \times \frac{8}{9} = \frac{49}{72} = 0{,}68$$

Pour toutes nos expérimentations nous essayons d'encourager les débats entre élèves. Rappelons que les échéances langagières sont les seuls permettant d'exprimer une interprétation de la probabilité. Dans le cas de ce problème en particulier, avec les proportions choisies nous cherchons, d'une manière ou d'une autre, à favoriser des débats. Les proportions des assemblages A et C donnent des probabilités presque égales. Ce qui pourrait dans un premier moment étonner les élèves vu que le troisième assemblage requiert le fonctionnement de trois pièces contre deux du premier. De sa part la probabilité de fonctionnement du deuxième reste la plus faible, contrairement à une approche intuitive qui indiquerait que vu qu'il ne requiert que le fonctionnement d'une seule pièce, il serait le plus probable de fonctionner.

Déroulement :

Référénciation par des ensembles finis

Une des différences entre la probabilité bayésienne et la fréquentiste consiste en que la première est une évaluation de la perception d'un individu et dans ce sens elle est subjective. De sa part, la fréquentiste décrit une caractéristique externe à lui, en tant qu'externe elle est couramment admise comme objective. Le procédé d'évaluation pourrait se décrire comme étant une introspection pour la bayésienne et une extrospection pour la fréquentiste. En effet, pour évaluer la probabilité bayésienne l'individu quantifie sa propre impression personnelle.

Pour évaluer une probabilité fréquentiste, il se place en tant qu'observateur d'un phénomène externe à lui.

Pour la probabilité bayé sienne, les sources d'évaluation numérique de cette perception peuvent être très variées, nous verrons lors de la troisième expérimentation un autre critère d'évaluation que celui que nous utilisons pour cette séance. Pour le problème de cette deuxième expérimentation, dont la nature de l'objet est une épreuve générique, l'évaluation numérique se fait par des ensembles de référence finis.

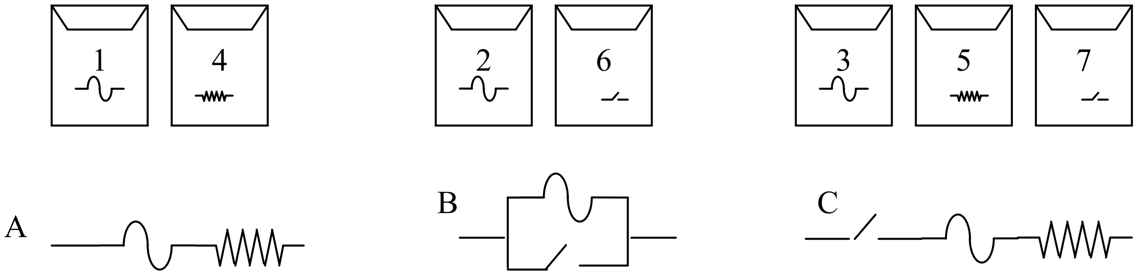
Ces sources d'évaluation, historiquement reconnues comme naturelles (*Chapitre I. La Logique de Port-Royal. Les* débuts de la quantification) lorsque elles sont présentées dans un problème et sans aucune autre source d'appréciation, il devient naturel que l'évaluation se réalise par la seule référence disponible. Dans les exercices de manuels (Chapitre II), toute autre source d'évaluation que les ensembles de référence est mise à l'écart par l'évocation du problème, l'élève n'ayant à disposition que l'ensemble de référence, il évalue son degré de certitude à partir de cette seule information.

La qualité de l'information

Néanmoins, dans le problème des circuits la situation est légèrement différente, les seules sources explicites d'information restent toujours les ensembles de référence finis (les composants dans les enveloppes). Mais le contexte, n'étant plus hypothétique ou évoqué comme dans les manuels, pourrait inciter les élèves à s'interroger sur la véracité de l'information concernant le fonctionnement des pièces, principalement parce que leur présence rend possible la vérification de leur fonctionnement. Ainsi, les élèves, dans leurs processus d'introspection, pourraient à juste titre vouloir intégrer leurs doutes sur le fonctionnement des pièces aux degrés de certitude.

Nous avons choisi pour cette séance de nous restreindre à l'utilisation des ensembles finis comme seule source d'évaluation. Afin d'éviter que le doute intervienne dans l'évaluation de la probabilité, nous avons considéré une première étape dont les élèves testent les composants à l'aide d'une pile et d'une ampoule. Ce test, en garantissant le maximum de qualité de l'information, écarterait les doutes comme sources d'évaluation. Lors de cette étape de vérification de fonctionnement, chaque groupe d'élèves teste une ou deux enveloppes, en indiquant sur sa couverture en papier la proportion de composants dèfectueux.

Visant toujours fournir le maximum de qualité possible, l'enseignant assemble les pièces à la vue des élèves, obtenant ainsi les trois circuits ouverts représentés sur la Figure 21 et reproduite ci-contre.



Puis, il pose la question du problème, montrant le circuit incomplet à fermer (ampoule plus pile), l'enseignant demande aux élèves de préciser lequel des trois assemblages on devrait prendre si l'on souhaite allumer l'ampoule. Les élèves, travaillant en groupes, devraient proposer une réponse et leurs raisons.

Quelques stratégies

Pour répondre à la question, les élèves peuvent se servir de plusieurs stratégies. L'une des plus complexes consiste à évaluer d'abord les probabilités des composants, puis, en interprétant la configuration de l'assemblage, reconnaître les opérations respectives permettant de déterminer la probabilité de fonctionnement de l'ensemble.

Par exemple, pour le premier assemblage, il s'agit de deux composants en série, le fonctionnement du système requiert le fonctionnement de ses deux éléments. Le calcul de leur probabilité de fonctionnement est relativement immédiat, il se fait par des ensembles de référence fini et indépendants, en projetant sur la pièce générique une caractéristique de l'ensemble auquel elle appartient (*Chapitre I. La référenciation*).

De leur part, les enveloppes différentes devraient induire la notion d'indépendance entre les extractions des deux pièces, ce qui conduirait à reconnaître dans la multiplication l'opération qui modélise cette caractéristique. De cette manière, l'évaluation de la probabilité de fonctionnement du premier assemblage requiert les probabilités de ses composants et l'opération produit.

Néanmoins, il nous semble improbable que les élèves s'engagent dans ce genre de démarche. Entre autres raisons, parce que ces notions, apprises au lycée, sont temporellement lointaines et utilisées dans des contextes assez différents de celui de cette séance (évoqués, sans prises de décision, sans interprétation, etc.). Cela rend ces notions difficiles à récupérer et à re-signifier dans une situation de relative nouveauté comme celle-ci.

En fait, il nous semble plus probable, par exemple pour le premier assemblage, que les élèves recourent à une démarche plus élémentaire, celle de l'énumération de toutes les combinaisons possibles d'assemblages. De cette énumération ils obtiendraient l'information suffisante au calcul de la probabilité, et cela à un coût relativement réduit en termes de temps et d'effort, en même temps que la démarche est sémantiquement facile à saisir.

Pour déterminer la probabilité de fonctionnement du premier assemblage, les élèves déclineraient donc tous les assemblages possibles en identifiant ceux dont tous les deux fonctionnent. De cette manière ils utiliseraient la formule de Laplace, non pas entre éléments, mais cette fois-ci entre assemblages possibles. Il s'agirait en fait d'une application du critère d'évaluation par des ensembles finis. Pour l'énumération exhaustive des cas favorables et possibles, il serait possible d'utiliser des listes de dénombrement ansin que des arbres ou des tableaux.

Pour ce qui concerne le troisième assemblage, l'utilisation de la même technique peut poser des difficultés dans le comptage, principalement pour ceux qui le font par des tableaux. Cette technique étant facile pour le premier devient complexe à cause des trois dimensions du troisième assemblage. Néanmoins, il est toujours possible de déterminer le nombre d'assemblages par une liste exhaustive ou par des techniques de dénombrement.

Pour sa part, le deuxième assemblage peut poser aussi des problèmes pour identifier le modèle mathématique correspondant à la configuration de ses composants ; une des erreurs les plus attendues consiste à déterminer la probabilité de fonctionnement de l'ensemble par la somme des probabilités des composants, sans effectuer la soustraction respective.

Quelques variables didactiques

Nous pourrions avoir minimisé la possibilité de cette erreur en choisissant d'autres proportions pour cet assemblage. Des proportions qui, lorsque cette erreur se produit, donnent des probabilités supérieures à l'unité, en attendant ainsi une réaction des élèves à cette valeur impossible de probabilité. Néanmoins, nous avons préfèré privilégier le débat qui pourrait provoquer le paradoxe qui représente sa faible probabilité de fonctionnement devant les autres assemblages. En ne requerant que le fonctionnement d'une seule de ses pièces son fonctionnement pourrait paraître plus probable que les circuits en série, dont aucune pièce ne doit être défectuese.

D'ailleurs, avec ces proportions nous cherchons non seulement à faciliter les conditions d'émergence des interprétations, mais aussi à créer les conditions qui justifient le calcul des probabilités. Un choix basé sur des arguments intuitifs indiquerait que l'assemblage le plus probable serait celui du circuit en parallèle (il suffit qu'un seul composant fonctionne). Si les élèves ne trouvent pas de raisons pour s'y opposer, il appartiendrait à l'enseignant de le faire, son point de vue s'imposerait non seulement par ses arguments mais aussi par son rôle en tant qu'enseignant. Les proportions choisies cherchent donc à faciliter aux élèves des arguments pour se prononcer contre les choix intuitifs.

De plus, lorsque ses proportions sont utilisées pour la négative (rejeter un argument intuitif) nous en attendons aussi pour la positive, en d'autres termes, si elles servent pour réfuter un argument intuitif, qu'elles servent aussi pour défendre un choix.. Les conflits entre approches devraient se résoudre par la précision que les calculs de probabilités apportent à la décision.

Les calculs de probabilité pour un même assemblage peuvent différer de groupe en groupe. Néanmoins, tout différend peut potentiellement s'assainir sans l'intervention de l'enseignant, en déclinant les cas favorables et possibles, en d'autres termes par une référenciation par des ensembles finis. D'ailleurs, il est légitime de s'interroger sur la plausibilité d'une référenciation par des ensembles infinis, la fréquence, avec laquelle on attend que les circuits fonctionnent lorsqu'on répète l'expérience. D'un point de vue pratique sa réalisation devient extrêmement coûteuse, nous admettons que cette difficulté de conception matérielle s'accompagne d'une difficulté d'utilisation comme argument pour évaluer ces probabilités. De cette manière, l'argument de la fréquence d'apparition, par son artificialité, ne serait très probablement pas évoqué par les élèves. Celle-ci est donc une autre variable didactique pour ce problème qui, additionnée à la présence matérielle des assemblages, cherche à fixer l'objet épreuve générique évalué par des ensembles finis.

Des difficultés

Nous nous attendons à que l'évaluation par des ensembles finis ne nécessite pas d'être explicitée dans ce problème. Plusieurs éléments jouent à sa faveur, premièrement, elle figure parmi les démarches les plus naturelles (Bernoulli, 1713; Nicole & Arnaud, 1662). Deuxièmement, elle est la plus fréquente dans les manuels (Chapitre II) et troisièmement, elle est la seule source immédiatement disponible. Dans ces conditions il est fortement probable que la classe tout entière s'en servira naturellement, tant pour les calculs que pour les vérifications.

La naturalité de la référenciation (évaluation par des ensembles finis) est fort utile, mais elle entraîne aussi ces difficultés. L'une de plus importantes consiste dans le risque que, à cause d'un usage exclusif dans les exercices, elle soit perçue par les élèves comme le seul moyen pour évaluer une probabilité bayésienne, induisant ainsi à installer une représentation objective de ce genre de probabilités. Une possible solution consiste à proposer des problèmes dont les ensembles de référence finis ne soient pas uniques, tel le cas de Camille (*Chapitre I. Quelques précisions sur les épreuves* génériques) dont l'évaluation de sa réussite à un examen peut se faire par plusieurs ensembles de référence finis, et de cette manière on peut obtenir plusieurs valeurs de probabilités pour une même épreuve. Malheureusement nous n'avons pas eu la possibilité d'expérimenter des tels genres de situations-problèmes. Néanmoins et même si la nature de l'objet est toute autre (hypothèse au lieu d'épreuve générique), nous aurons l'occasion d'expérimenter, lors de la troisième séance, l'évaluation d'une probabilité dont elle n'est pas immédiate, comme c'est le cas pour cette deuxième expérimentation.

Des notions à institutionnaliser

De la même manière que pour la première expérimentation, celle-ci a été conçue pour qu'elle profite, non seulement à notre recherche, mais aussi au programme de l'enseignant. En particulier cette expérimentation répond à des objectifs institutionnels liés aux axiomes de Kolmogorov. Tous les assemblages abordent des notions de dénombrements en cas favorables et possibles. De plus, le premier et le troisième des assemblages permettent le traitement de la notion d'indépendance stochastique et le deuxième celle de probabilité de l'union de deux événements.

Ces notions ont été déjà traitées pour ces élèves lors de leur passage au lycée, mais étant les élèves d'origines institutionnelles diveres, l'enseignant nous a proposé de les récupérer dans une de nos situations-problèmes. En nous basant sur cette proposition nous avons conçu cette situation qui répond, dans notre approche, à un intérêt d'observer les possibilités de traitement de l'interprétation de la probabilité bayésienne dans sa modalité épreuve générique.

Dans cette deuxième expérimentation et de la même manière que pour la première, nous avons partagé avec l'enseignant quelques impressions sous la forme d'un possible déroulement. Ces échanges ont eu, comme point de départ, quelques notes que nous lui avions fait parvenir. Néanmoins, et à la différence de la première expérimentation, ces notes se sont cernées à expliquer quelques-uns de nos choix, en particulier nous lui avons décrit :

- la fonction des variables dans le problème (présence matérielle des objets, les proportions dans les enveloppes, etc.)
- l'interprétation de la probabilité associée au problème.
- la nécessité de favoriser les débats.
- l'importance d'institutionnaliser l'interprétation de la probabilité.
- la place de la probabilité dans une approche décisionnelle.

Pour cette deuxième expérimentation, les impressions que nous avons partagées avec l'enseignant sous la forme d'un document brouillon d'analyse *a priori*, ont été d'une modalité différente de celle de la première expérimentation. Lors de la première séance et même dans notre analyse *a posteriori*, nous avons eu l'impression que l'enseignant s'était attaché au possible déroulement décrit dans notre analyse *a priori* et que, dans cet attachement il avait, soit restreint sa liberté de choix, soit écartée certaines étapes afin de rester proche du déroulement *a priori*.

Pour cette deuxième expérimentation nous avons préfèré de ne pas décliner l'analyse en étapes comme lors de la première, mais plutôt lui faire savoir nos choix en termes de variables et des enjeux pour notre recherche.

Donc, pour dette deuxième situation, et afin de restituer à l'enseignant son espace de choix dans la classe, nous avons décidé de nous limiter à partager les items ci-dessus énumérés.

La probabilité conditionnelle

Nous avons prévu, si le temps le permettait, d'introduire une variante au problème nous permettant d'approfondir la notion bayésienne de la probabilité. A la de suite du problème proposé et, une fois la question du choix du circuit répondue, l'enseignant prendrait un des assemblages et en le démontant, il choisirait une des pièces et la testerait en se servant de la pile et de l'ampoule, puis en replaçant la pièce dans le circuit, il demanderait encore une fois d'un testeur (pile plus ampoule) lequel des trois assemblages les élèves choisiraient, cette fois-ci en prenant compte de la nouvelle information.

Une autre variante consiste à partager cette nouvelle information avec un groupe de la classe. De cette manière, il y aurait des groupes évaluant des probabilités sur les mêmes objets de manière différente.

Cette variante permet, d'une part et d'une manière simple, de mettre en évidence le caractère non fixe de la probabilité bayésienne et de sa dépendance à l'information disponible. D'autre part, elle permet d'introduire le symbolisme de la probabilité conditionnelle. Par la suite nous analyserons l'expérimentation de cette deuxième situation-problème.

Analyse *a posteriori*

La salle et son équipement

La Figure 24 est une représentation de la distribution des vingt élèves (cinq groupes : G1 à G5) et du matériel utilisé pour la séance. Nous avons placé quatre dictaphones sur les tables des élèves (1 à 4), un vidéo projecteur (5) et une caméra (6) au fond de la salle.

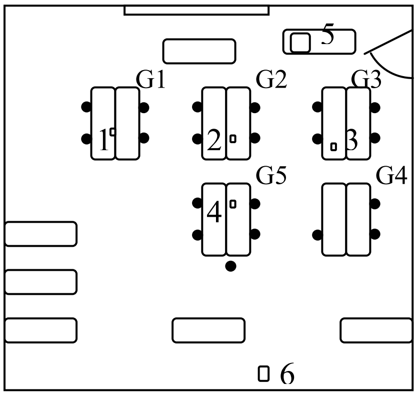
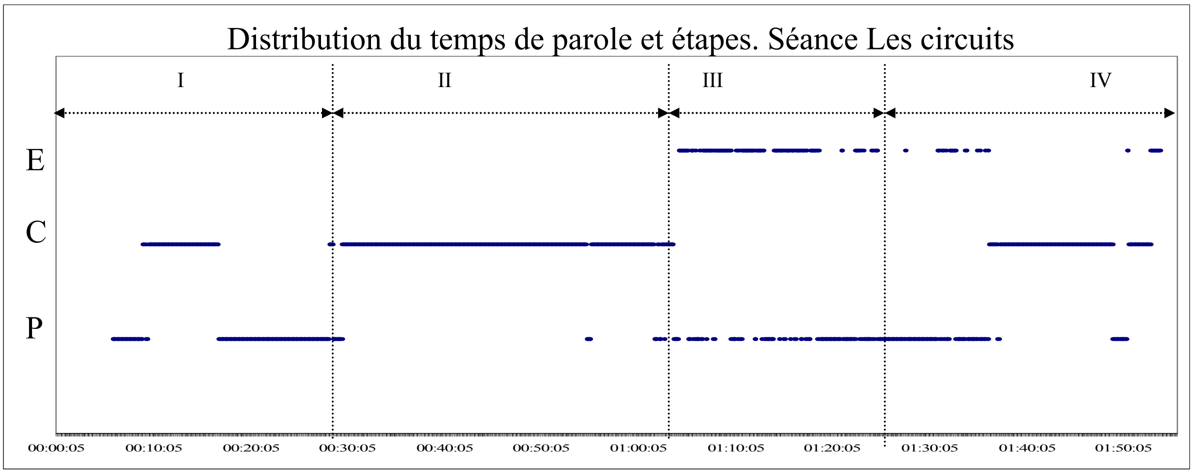


Figure 24

Le Graphique 43 est une représentation de la séance qui s'est déroulée le 13 novembre 2006 où nous avons expérimenté cette deuxième situation-problème. Pour ce graphique nous avons pris en compte les mêmes critères que pour la première séance. A chaque fois qu'un membre d'un de ces groupes émettait un message, nous notions un point sur l'échelle chronologique. A chaque intervalle de cinq secondes de parole nous ajoutions un autre point. En annexe nous avons placé le décryptage des échanges enregistrés:

- Groupe P : L'émetteur du message est l'enseignant et il s'adresse à toute la classe.
- Groupe E : L'émetteur du message est un élève et il s'adresse à toute la classe.
- Groupe C : L'émetteur (enseignant ou élève) s'adresse à un membre de la classe ou à un groupe. Généralement, ce groupe était constitué par des élèves voisins travaillant ensemble pour résoudre le problème.

De la même manière que lors de la première séance, nous nous servirons de cette représentation pour notre analyse *a posteriori*.



Graphique 43

Quatre étapes résultent de ce graphique, la première concerne la présentation du problème, y compris le test des pièces par les élèves, la deuxième l'élaboration de propositions des élèves ; elle se caractérise par des débats au sein de chaque groupe, la troisième où les élèves présentent au reste de la classe leurs conclusions, et la dernière où l'enseignant réalise l'institutionnalisation de la séance. Nous commençons l'analyse pour la première étape.

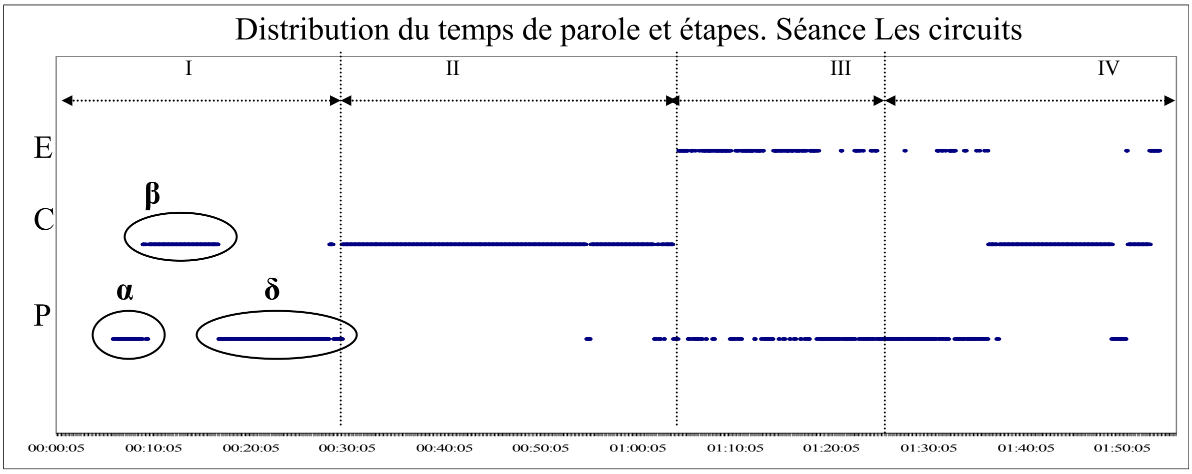
Pour cette analyse a posteriori nous nous servons des données provenant des sources suivantes :

- Décryptages des quatre dictaphones (G1, G2, G3 et G5)
- Vidéo de la séance
- Présentations en transparents de chaque groupe (G1 à G5)

  Feuilles de recherches recueillies à la fin de la séance (G1, G2, G4 et G5).

Présentation du problème

A la minute six de nos enregistrements (I, Graphique 44), l'enseignant présente la première partie du problème, celle qui concerne l'étape de test des pièces (α). A la différence de la séance précédente, les élèves se trouvent assis en groupes de trois à cinq ; cette organisation avait été concertée lors de nos entretiens ayant précédé la séance.



Graphique 44

Pendant la présentation, l'enseignant distribue les enveloppes en papier contenant les pièces et les testeurs (une pile et une ampoule) aux élèves afin qu'ils réalisent les tests[18]. Cette tâche commence à s'effectuer lentement (β). Des quatre groupes où nous avons placés les dictaphones, nous ne retenons rien de significatif de cette étape de test, probablement que l'absence de réflexions et de débat peut s'expliquer par la méconnaissance de la part des élèves des objectifs de cette tâche. En effet, l'enseignant sollicite de tester les composants sans expliquer le but poursuivi.

L'absence d'indications justifiant l'activité sollicitée pourrait trouver plusieurs raisons. Nous en retenons deux. L'une concerne les aspects organisationnels, l'autre est liée aux pratiques profesionnelles de l'enseignant, toutes les deux liées.

L'explication organisationnelle traite sur les possibles dérangements qui pourrait avoir occasionné l'explicitation du projet sans que l'information nécessaire à sa résolution soit encore à disposition des élèves. Par exemple, ayant terminé de tester leurs pièces, quelques élèves pourraient réclamer aux autres de se dépêcher pour ainsi résoudre le problème. De plus l'information concernant les proportions pourrait circuler de manière non vérifiée et probablement se dénaturer. D'ailleurs, une description complète du problème entraînerait une impression d'inactivité des élèves, ce qui pourrait inquiéter l'enseignant vis-à-vis de la nécessité de faire bouger la situation (Robert, 2001). Bref, des facteurs liés à des aspects

[18] A la place de fusibles nous avons utilisé des connecteurs, ces derniers se sont révélés plus facile à abîmer que les fusibles. D'ailleurs, pour l'enveloppe 6 (interrupteurs) du test réalisé par les élèves il en résulte trois abîmés sur neuf à la place de deux sur neuf prévus *a priori*

organisationnels pourraient expliquer le choix de l'enseignant de présenter le problème après le test des pièces.

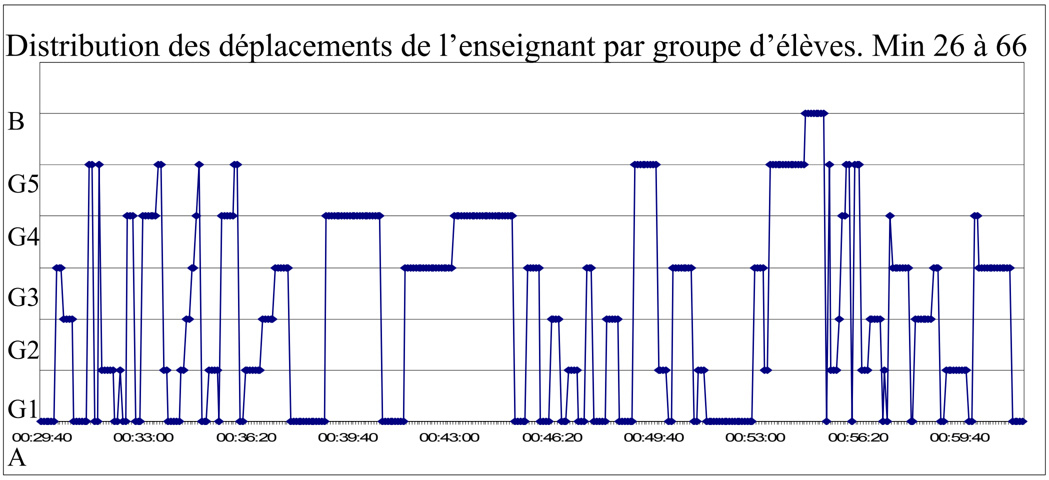
Une autre explication nous semble possible, elle est liée aux habitudes du métier, habitudes qui correspondraient d'une certaine manière à une habitude courante des manuels. En fait, la structure sous-jacente à cette présentation est assez proche de celle des exercices que nous avons analysées. Cette structure, que malheureusement nous n'avons pas statistiquement analysé, se caractériserait par, d'abord une description du contexte, puis une présentation de l'information nécessaire à sa résolution et finalement la (les) question(s) à résoudre. Il semblerait que cette tendance dans les manuels se reproduirait dans l'explication de l'enseignant. Cette structure pourrait donc être, non seulement exclusive aux manuels mais devenir une tendance dans l'enseignement des mathématiques.

Suite au test du fonctionnement des pièces et à l'aide du rétroprojecteur, l'enseignant présente les résultats sur l'écran de la salle. Puis il commence à monter les circuits. Après huit minutes, les assemblages prêts, l'enseignant pose la question du problème :

23 *Bon alors j'ai entendu j'ai entendu des questions j'ai entendu pourquoi est-ce qu'on fait ça en maths et voilà la question la question qu'on va se poser c'est la suivante je voudrais allumer l'ampoule du testeur avec un de ces trois circuits avec un de ces trois circuits peu importe le voltage peu importe l'intensité du courant je vous demande quel est de ces trois circuits quel est celui que vous me conseilleriez d'utiliser*

Elaboration de propositions

A la minute vingt-six, les groupes commencent à discuter les solutions possibles (II, Graphique 44). Dans cette étape de quarante minutes, l'enseignant n'intervient que rarement. Il visite constamment les différents groupes tout en restant à l'écart des débats. Le Graphique 45 montre la distribution de ses déplacements pour les respectifs groupes (G1 à G5). Pendant cette période, il consacre quelques minutes aussi à des tâches administratives ou à la révision de son organisation pour la séance (A). Un des groupes (G4, sans dictaphone) élabore des idées qui attirent son attention ; il nous le fait savoir en B.



Graphique 45

De leur côté, les élèves s'habituent progressivement à la présence des dictaphones sur leurs tables. Certains groupes sont moins indifférents que d'autres à être enregistrés.

La classe s'engage très rapidement à déterminer les probabilités de fonctionnement des montages. A la différence de la première séance, personne n'interpelle le procédé de décision basé sur une probabilité (min 36. « **Pour moi c'est celui qu'a le plus de probabilité** »). Même l'enseignant nous le signale à la fin de la séance :

> *« (…) j'ai eu l'occasion de parcourir un peu la production de chaque équipe mais d'une part disons que la probabilité est sortie naturellement pour prendre la décision je m'attendais à un petit peu de conflit et non ils ont pris ils l'ont utilisé (…) »*

Néanmoins, derrière cette apparente acceptation de l'approche indéterministe encadrant la solution, la mise en place d'un contrat didactique pourrait s'occulter, ce qui ferait que les élèves, en s'apercevant des tâches mathématiques à effectuer, les réalisent sans s'interroger sur les aspects argumentatifs de la solution. Toutefois, cette acceptation ne fait pas l'unanimité, dans l'étape III, lorsque l'enseignant cherche à faire débattre les élèves sur le critère décisionnel, l'un d'entre eux semble garder une position déterministe (min 85 : « **Mais on ne peut pas savoir** »), nous y revendrions plus tard lors de l'analyse de l'étape III de cette séance.

Pendant cette étape de travail en équipe, les débats se nourrissent non seulement des propositions au sein du groupe, nous avons constaté quelques échanges entre les groupes et même des idées empruntées à des groupes voisins. Ces idées exogènes sont en général discutées mais rarement assimilées, principalement à cause de la distance qui les sépare des élaborations endogènes. En général, les idées exogènes concernent des aspects calculatoires, liés à la modélisation du problème. Pour son assimilation, le groupe les analyse sémantiquement, si sur cette dimension ces concepts ne parviennent pas à être justifiés, ils ne sont pas retenus. De cette manière, la sémantique des objets fonctionne comme un filtre pour les idées exogènes, l'idée est retenue si les élèves comprennent la relation entre l'objet mathématique et la caractéristique du problème associée. Ce mécanisme s'est observé aussi sur quelques idées endogènes (min 49, « **Tu as fait comment ça expliques (…) Attends explique pourquoi tu as fait comme ça**»). Néanmoins, toutes les idées (endogènes, principalement) ne sont pas soumises à ce genre d'analyse, quelques-unes étant admises sans s'interroger sur leur signifiés.

En ce qui concerne les assemblages, ils sont bien compris pour les élèves. Il est connu pour cette classe de BTS qu'un circuit en série requiert le fonctionnement de tous ses composants et qu'un circuit en parallèle n'a besoin que d'un seul. La principale difficulté concerne la relation entre l'interprétation du problème et sa modélisation, plus précisément la manière dont la probabilité est calculée en fonction du type d'assemblage. Nous avons identifié plusieurs schémas de la relation entre l'interprétation du problème et sa modélisation, nous les présentons ci-après.

Schéma I

Ce schéma correspond à des cas où nous constatons une compréhension des enjeux de la situation (ou d'une partie) accompagnée d'une impossibilité à trouver une modélisation adéquate.

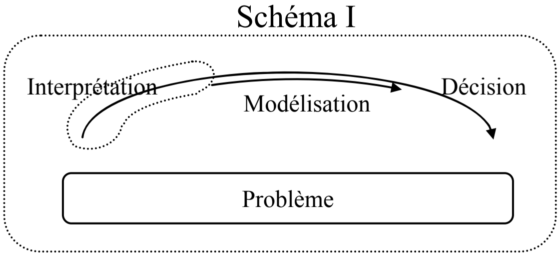


Figure 25

Par exemple, le Groupe 2 sait qu'une seule pièce suffit pour que le circuit en parallèle fonctionne, et que la probabilité de fonctionnement de l'ensemble est supérieure à celle de ses composants, contrairement aux circuits en série dont la probabilité de fonctionnement de l'ensemble est inférieure à celles de ses parts. La reconnaissance de ces relations d'inégalité place les élèves dans ce que nous pourrions appeler une phase de pré modélisation, dans laquelle sont reconnues des caractéristiques relationnelles (majeur, mineur que…) mais qui n'arrivent pas à une formalisation mathématique. Pour ce groupe, l'évaluation de cette probabilité reste inaccessible, néanmoins les élèves décident de retenir les conclusions de cette étape en présentant la cote inférieure de la probabilité de fonctionnement du circuit en parallèle, minutes 34 et 50) :

34 **Pour le deuxième cas au minimum tu as deux chances sur huit minimum** (…) **j'explique le deuxième cas au minimum tu as deux chances sur huit** (…)

Les difficultés du Groupe 2 à trouver la valeur de la probabilité pourraient s'expliquer par la stratégie retenue pour ce groupe lors des calculs des probabilités. Nous revenons sur les deux principaux types de stratégies possibles pour ce problème, nous en dénommerons une *directe*, l'autre *indirecte*.

Stratégie directe

Elle consiste à évaluer une probabilité par un ensemble de référence, dans ce cas, un ensemble fini. Par exemple, les probabilités des composants sont évaluées de manière directe, les élèves utilisent leurs tests pour déterminer le rapport entre les cas favorables et les cas possibles.

L'application de cette technique au calcul des probabilités des assemblages implique la détermination, par des moyens divers, des cardinaux respectifs (cas favorables et possibles). Cette stratégie est la plus facilement saisissable en termes sémantiques.

Stratégie indirecte

Elle consiste à déterminer une probabilité à partir d'autres probabilités, sans une utilisation d'un ensemble de référence, dans ce sens elle est indirecte. Dans ce cas, les probabilités des assemblages s'obtiendraient à partir des probabilités des composants et de la configuration du montage. Cette stratégie, plus complexe que la précédente, implique un travail de suivi sémantique en même temps que des transformations tant sur le plan arithmétique que sur les objets. En effet, la dimension sémantique reste inaltérée (degré de certitude), les objets sur lesquels portent les probabilités changent (des composants vers l'assemblage) du même que les valeurs pour les opérations concernées (produit, adition, etc.). La gestion de tous ces aspects implique un travail plus complexe que celui de la stratégie directe.

Le Groupe 2, au début de cette étape, se sert des deux stratégies pour le calcul de la probabilité du premier montage (A, circuit en série). Un élève la détermine par la *directe*, un autre par l'*indirecte*, tous les deux obtiennent la même valeur. Lors de la confrontation des deux résultats ils concluent que les stratégies sont équivalentes. Par la suite, et pour des raisons diverses (moyen plus économique, ou plus proche des attentes en classe de mathématiques, etc.) les élèves du groupe retiennent la stratégie indirecte. Celle-ci leur permet de réussir leurs calculs pour les deux circuits en série, mais elle devient trop complexe pour le montage B en parallèle.

En effet, pour le circuit en parallèle la technique indirecte est, sans le secours des ensembles de référence, plus compliquée que pour les circuits en série. Ce groupe qui reste attaché à la stratégie indirecte ne parvient pas à revenir sur la directe pour, en se référenciant sur les cas possibles et favorables déterminer la probabilité de fonctionnement du deuxième montage.

Ce schéma représente donc les initiatives sémantiquement correctes mais tronquées par l'impossibilité de lui associer une modélisation satisfaisante. Le Groupe 2 nous a permis d'illustrer ce schéma, néanmoins, si nous prenons en compte leurs productions pour déterminer la probabilité de fonctionnement de l'assemblage A, ils se rapporteraient à un autre schéma. En fait, les schémas retenus ne constituent pas une bijection avec les groupes, ils représentent les différentes relations que les groupes établissent entre les aspects sémantiques et numériques des probabilités calculées, de cette manière un même groupe se correspond avec plusieurs schémas.

Schéma II

A cet autre schéma correspond le diagramme de la Figure 26, il représente un travail exclusif sur le plan de la modélisation mathématique sans support sémantique. Dans cette séance nous l'avons observé lorsque les élèves empruntent un modèle à un autre groupe. Son assimilation et utilisation au sein du groupe dépendent, entre autres, de la distance qui sépare ces nouvelles idées de celles du groupe. Un autre facteur pour cette assimilation semble être le rôle occupé par chacun des élèves au sein du groupe. Il semblerait que les idées portent une pondération en fonction de l'influence des élèves (plus ou moins bons en mathématiques, etc.) (« **Je pense qu'il faut additionner parce que j'ai tort par rapport à toi en maths** », « « **Et Jérôme là il a fait quoi** »).

Schéma II

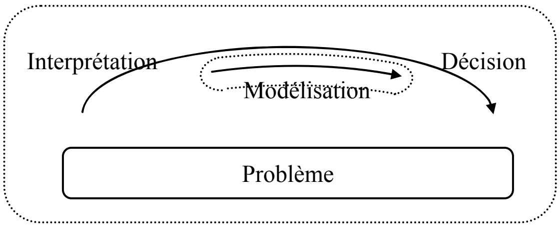


Figure 26

Nous retrouvons un exemple de ce schéma dans le Groupe 2, qui semble lui avoir emprunté des idées à deux reprises, et à des groupes différents. Nous décrirons l'emprunt de la solution au montage B :

63 **Et le deuxième quarante six pour cent on est bien d'accord**
De toute façon c'est le dernier
Mais non attends le deuxième c'est bien quarante six pour cent

Si pour évaluer la probabilité de fonctionnement du circuit B, nous prenons en compte une référenciation par la proportion de cas favorables et possibles, la valeur numérique de la probabilité est en effet celle indiquée par cet élève (« **Quarante six pour cent** »). Néanmoins, nous n'avons pas trouvé de traces des étapes lui permettant d'arriver à une telle valeur, ni dans les feuilles de recherche ni dans le décryptage, ni dans la présentation à la classe non plus, ce qui nous mène à penser que cette valeur a été empruntée à un autre groupe, probablement aux groupe 4 et 5, les seuls à avoir fourni des traces de la construction de cette réponse. Toutefois, cette valeur n'est pas retenue par le groupe qui finit par l'écarter vu l'impossibilité de la défendre.

Ce schéma n'est pas exclusif du Groupe 2, d'autres tels que le Groupe 1, l'ont essayé (min 57 : « **Et Jérôme là il a fait quoi** ») mais sans réussir à obtenir l'information cherchée (« **Bah il dit qu'il faut voir** »).

Schéma III

Ce troisième schéma représente la réunion des deux précédents (Figure 27), mais de manière non articulée. D'une part, les élèves donnent des indices de compréhension des enjeux du problème, et d'autre part, la modélisation tentée ne correspond pas à la configuration du montage. Dans certains cas, cette non correspondance entre modèle et réalité est reconnue par les élèves, dans d'autres le modèle retenu est pris comme représentatif du montage en question.

Schéma III

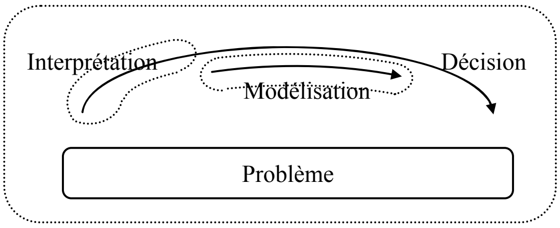


Figure 27

Nous avons observé ce schéma dans les deux stratégies possibles, tant la directe que l'indirecte. Pour ce qui concerne la directe, la source de la non correspondance entre modèle et réalité provient des erreurs dans l'énumération de cas possibles et favorables (Présentations Groupe 1 et 3. Figure 28). Le manque d'une technique exhaustive d'énumération semble être à la source de l'erreur.

Erreurs dans la stratégie directe

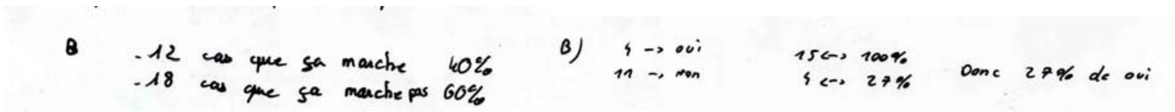


Figure 28

De sa part, pour la stratégie indirecte, les raisons sont diverses, nous avons observé deux types. Le premier consiste en une non maîtrise des règles de transformation dans le registre numérique (Présentation Groupe 2 Figure 29). Dû à une erreur dans le calcul des aditions de fractions, le modèle perd sa représentativité de la réalité (le modèle est de toute manière incomplet, il ne considère pas les doublons dans le comptage).

Erreur dans la stratégie indirecte (règles de transformation)

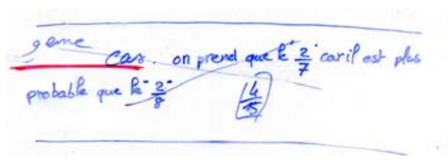


Figure 29

Le deuxième concerne des incomplétudes. Ces erreurs correspondent à la non prise en compte des doublons dans l'addition des probabilités pour le montage B (Présentation Groupe 3, Figure 30), erreur aussi présente dans l'exemple précédent.

Erreur dans la stratégie indirecte (incomplétude du modèle)

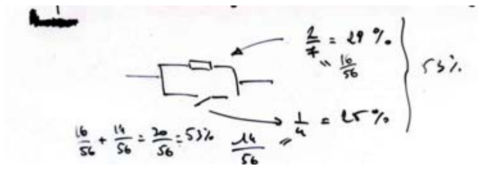


Figure 30

Le troisième type concerne des erreurs par la non correspondance entre l'opération choisie et la configuration du montage. Dans ces cas, l'opération ne modélise pas les cardinaux des ensembles de référence (Feuille de recherche Groupe 2) :

Erreur dans la stratégie indirecte (non correspondance des opérations)

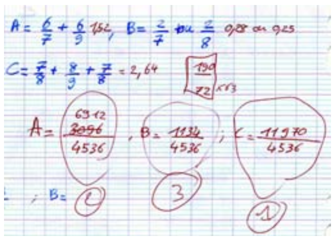


Figure 31

Schéma IV

Ce schéma correspond aux cas dont le modèle mathématique retenu par les élèves représente les caractéristiques du problème. Nous avons trouvé des exemples de ce schéma dans les deux types de stratégies du calcul des probabilités, directe et indirecte.

Schéma IV

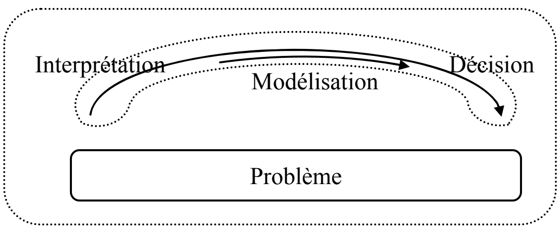
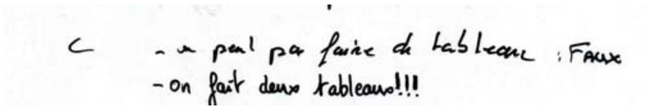


Figure 32

Pour ce qui concerne la directe, le Groupe 1 est le seul à avoir trouvé un moyen pour déterminer les cardinaux pour le troisième circuit. Leur stratégie consiste en décomposer l'énumération des cas en deux étapes (Présentation Groupe 1).

Calcul de probabilité pour le montage C par stratégie directe



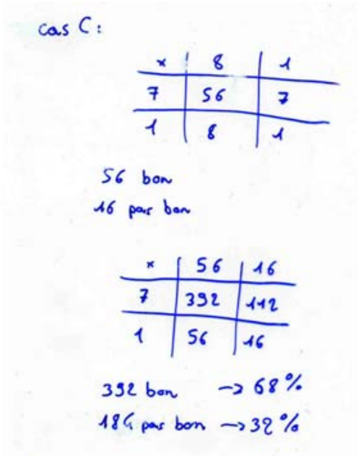


Figure 33

Pour ce qui concerne une utilisation de la stratégie indirecte pour trouver les probabilités des trois montages, un exemple est constitué du Groupe 4 où, dans sa présentation un élève précise tant les calculs que les expressions dans le registre symbolique (Présentation Groupe 4) ayant conduit à la détermination des probabilités :

Calcul des probabilités par stratégie indirecte

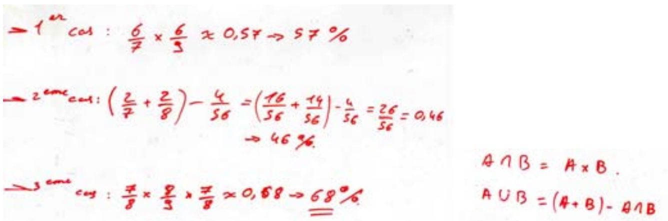


Figure 34

Des élèves du Groupe 4 nous ne disposons d'aucune trace de leurs débats (seul groupe sans dictaphone). Néanmoins, à partir des feuilles de recherche, il nous semble que cette formulation représente un seul élève, les autres, toujours en se servant de la stratégie indirecte, auraient trouvé des difficultés à déterminer la probabilité de fonctionnement du montage B. Nous ignorons de plus de quelle manière cet élève établit les liens entre la formulation (symbolique et numérique) et les caractéristiques du problème (sémantique).

D'ailleurs, la présentation des résultats du Groupe 4 (minute 72) suscite des réactions dans la classe :

- des élèves qui demandent l'origine de la formule, l'élève répond l'avoir appris au lycée.

- d’autres, même s’ils l’ont apprise, reconnaissent ne pas y avoir pensé.
- Et finalement d’autres nient une possible utilité de cette formule pour résoudre le problème.

Ces réactions semblent mettre en évidence l’importance de la relation entre les objets mathématiques et ceux du contexte, en d’autres termes la sémantique des objets mathématiques. Sans la compréhension de leur signification, les élèves tendent à ne pas les intégrer. La nécessité de correspondance entre l’interprétation du problème et les objets mathématiques est une question récurrente dans les échanges des élèves. Lorsqu’un élève propose un nouvel outil, les autres demandent des précisions (« **il faut que tu expliques pourquoi** »).

D’ailleurs, les explications que les élèves fournissent nous restent toujours difficiles à déceler, probablement dû aux difficultés que trouvent les élèves pour prendre une position réflexive par rapport à eux-mêmes (métaconnaissances).

Schéma V

Nous avons trouvé un dernier schéma pour cette séance. Celui-ci consiste en l’extrapolation d’une modélisation vers une autre plus complexe, extrapolation qui se base sur la similarité des structures des montages. Par exemple, le modèle multiplicatif du montage A (circuit en série de deux composants), est couramment extrapolé vers le troisième montage (C, circuit en série de trois composants) sous l’argument de leur similarité. La Figure 35 représente cette démarche, qu’est en général prise par les élèves de manière presque immédiate.

Schéma V

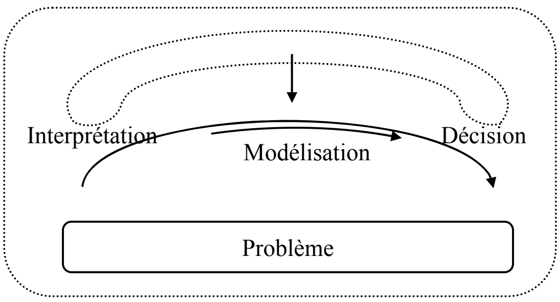


Figure 35

A la minute soixante trois, les différents groupes commencent à présenter leurs conclusions.

Les présentations des groupes

Au début de l'étape II l'enseignant avait distribué des transparents, un pour groupe. Ces transparents sont utilisés pour les élèves pour projeter sur l'écran de la salle leurs résumés) Les présentations se succèdent dès la minute soixante trois (Groupe 3) jusqu'à la quatre vingt (Groupe G5).

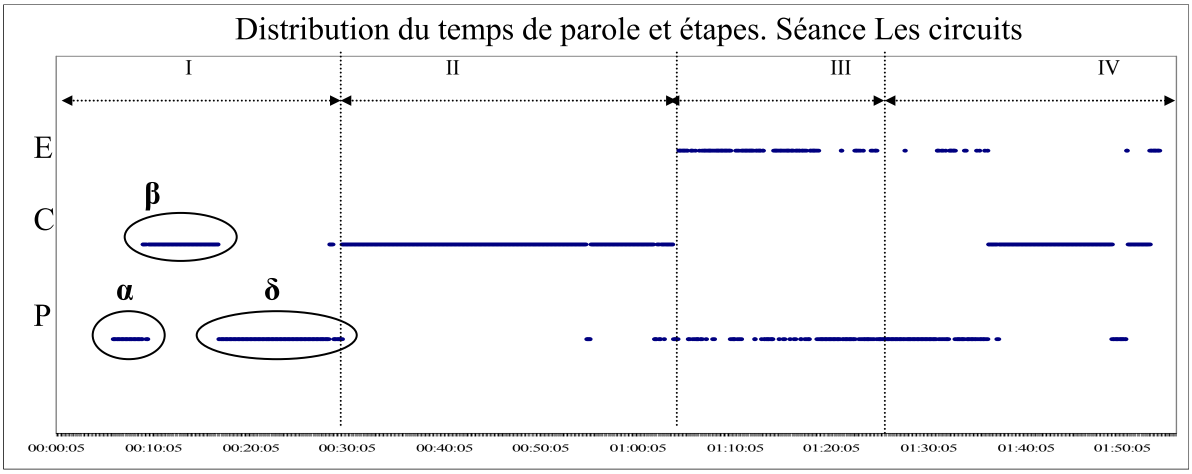


Figure 36

Dans leurs présentations, les élèves se focalisent sur les méthodes employées pour déterminer les probabilités. Le critère décisionnel, basé sur la probabilité maximale, reste relégué dans ces expositions et même il l'a été dans les débats ayant précédé les présentations. Un seul des groupes précise le critère utilisé (minute 75, G5, « **on a pris le cas qui a la plus forte probabilité de fonctionner, bah par rapport aux autres** »). Les autres exposent directement les stratégies du calcul des probabilités.

Les stratégies exposées correspondent les unes à la directe, les autres à l'indirecte. Le Groupe 3 développe son explication à partir de la première, leurs valeurs de probabilité n'étant pas correctes, les erreurs sont dues à des dénombrements incomplets. Pour sa part, le Groupe 2 réussit par l'indirecte à "bien"[19] déterminer les probabilités des deux montages en série. Le troisième groupe à exposer (G1), s'est servi pour les trois montages de la stratégie directe pour trouver les valeurs de probabilité. Dans leur transparent ils montrent les cardinaux trouvés pour les trois assemblages. En bas, ils ont ajouté comme illustration de la

[19] Dans l'approche bayésienne il n'existerait pas une valeur correcte pour évaluer un degré de certitude, ce qui existait pour cette approche se sont des raisons pour l'évaluer. Dans cette séance, pour des ensembles de référence finis.

méthode, le tableau utilisé pour déterminer les cas favorables et possibles pour le montage A. C'est le seul groupe à avoir tenté de résoudre le montage en parallèle par la stratégie directe et à avoir trouvé par la même stratégie la probabilité pour le troisième montage (en série). Soulignons que jusqu'à présent personne ne sait si les valeurs présentées sont correctes ou non, l'enseignant garde toujours la même position que dans l'étape précédente (*Elaboration de propositions*), il ne fait que modérer les interventions des élèves.

Le contrat didactique

La stratégie directe semble, à ces élèves, ne pas répondre aux attentes, personnelles et/ou institutionnelles. Leur présentation commence avec un regret (min 69, « **on n'a pas fait grand chose on n'a pas trouvé une méthode à la fin on a fait des tableaux pour le B** »). Il semblerait que l'attention des élèves s'oriente exclusivement à résoudre le problème mais aussi à le faire par des moyens spécifiques, ceux qui se caractérisent par un formalisme mathématique. Cette condition de formalité mathématique de la solution est en tout cas implicite et ferait partie du contrat didactique de la classe de mathématiques, nulle part l'enseignant n'a donné des consignes de ce genre dans cette séance. De cette manière, le contrat didactique semble guider les choix des élèves, qui privilégieraient certain genre de solutions (formelles) en détriment d'autres (non formelles).

La nécessité de répondre aux attentes concernant le formalisme mathématique pourrait expliquer les difficultés de quelques groupes à déterminer la probabilité de fonctionnement du montage en parallèle. En général les élèves arrivent à trouver une représentation "attendue" pour les montages en série, le produit des probabilités. Mais ils ne parviennent pas à en proposer une pour le montage en parallèle. Il semblerait que, motivés par le but d'une solution formelle, ils écartent d'autres voies qui bien que tout à fait valables, ne répondent pas aux attentes du contrat didactique. De cette manière, ils perdent la possibilité d'utiliser des méthodes simples pour résoudre le problème.

D'ailleurs, la reconnaissance de la part des élèves des tâches attendues pourrait expliquer l'absence de débat concernant le critère décisionnel. Ayant reconnu ce qu'il faut faire, ils s'y investiraient sans trop réfléchir sur l'enchaînement argumentatif du problème. Ce facteur pourrait donc expliquer l'absence d'indices dans les débats concernant le critère décisionnel.

Une autre explication à l'absence de débat sur le critère décisionnel pourrait se trouver dans la combinaison de deux caractéristiques de ce problème, des caractéristiques assez différentes à celles de la première situation-problème. Ces caractéristiques sont d'une part la nature des objets sur lesquels portent les probabilités, de l'autre, la complexité du critère décisionnel.

En effet, le critère décisionnel pour cette deuxième situation est plus simple que celui pour la première situation. L'un se base sur la probabilité, l'autre sur l'espérance mathématique. De plus, pour le deuxième problème la probabilité porte sur les objets en question, les montages. Pour l'autre, sur ce qui arriverait lors d'une infinité de tirages. Il semble plus naturel d'accepter les principes argumentatifs de ce problème que ceux du premier. De cette manière, le critère pour ce deuxième problème s'admettrait naturellement, sans nécessité de discussion.

A la fin de la présentation du Groupe 1 l'enseignant commence à expliciter quelques éléments qu'il reprendra après en l'étape d'institutionnalisation. Il demande au Groupe 1 (stratégie directe) de préciser la méthode utilisée pour trouver les cardinaux des ensembles pour le montage en parallèle. Le groupe ne donne pas la réponse tout de suite, pour cela, les élèves se réunissent autour de leurs tables. La réponse viendra après la présentation du Groupe 5. Nous verrons, suite aux présentations des élèves, que l'enseignant retiendra la stratégie directe du calcul des probabilités. Cette stratégie, en considérant l'état d'avancement des élèves, semblerait à l'enseignant plus convenable à institutionnaliser que l'indirecte. Surtout parce qu'il vise l'institutionnalisation aussi du signifié de la probabilité en tant que degré de certitude. Nous reviendrions plus tard sur ce sujet.

Le Groupe 4 suit le Groupe 1. Cette nouvelle présentation se place dans une position diamétralement opposée à la précédente, ces élèves font l'exposé le plus formel de la séance. A partir de la stratégie indirecte ce groupe présente les trois calculs dans le registre numérique et les respectives expressions génériques dans le symbolique, leurs calculs étant tous corrects. Néanmoins, les réactions des élèves sont bien différentes à celles de la présentation précédente (stratégie directe). La présentation du Groupe 1 était sémantiquement plus saisissable pour les élèves, ses calculs se basaient sur une référenciation par des ensembles finis. Celle du Groupe 4 s'effectue à partir des probabilités des composants.

Les élèves interrogent tout de suite le représentant de ce groupe. Les questions ne concernent pas les circuits en série, en général la stratégie indirecte pour ces montages est comprise. Nous pourrions situer la classe entière dans les schémas IV et V pour les montages en série A et C respectivement. Néanmoins face à la présentation formelle du calcul de la probabilité pour le circuit en parallèle, la clase se trouve dans le schéma III, sans possibilité d'établir des liens entre les caractéristiques de ce montage et l'expression mathématique le représentant. Même les élèves du Groupe 1, étant arrivés à trouver les cardinaux des montages favorables et possibles pour le montage en parallèle, ne parviennent pas à reconnaître les signifiés des termes de la formule du Groupe 4.

L'écart entre la classe et la présentation de ce groupe fait que quelques élèves demandent des explications sur l'origine de la formule utilisée. Cette tâche demandée à l'élève du Groupe 4 n'est pas évidente. Il doit confronter des difficultés de différents types, d'une part émotionnelles : le stress de se trouver devant toute la classe, un étranger en train de l'enregistrer, etc. ; d'autre part métacognitives : pour répondre à la question il doit réfléchir sur ses propres actions et pouvoir les transmettre.

Formulation en registre symbolique pour le montage en parallèle (Groupe 4)

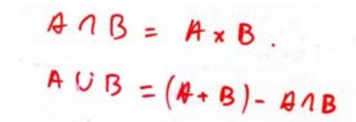


Figure 37

L'élève, dans sa réponse, résout ces difficultés en se centrant sur les liens entre le problème et sa modélisation. Il signale les types d'événements composés (A∩B et A∪B) repérés dans le problème (min 74, « **C'est la formule qu'on a apprise de a et b et d'a ou b** »). Cet élève reconnaît le genre de situations auquel chacun de ces types d'événement représente, mais ce n'est pas le cas pour la classe qui semble avoir attendu un autre genre d'explications (« **C'est une formule automatique** », « **Mais çà sert à rien cette formule** »).

L'élève a abordé sa réponse très probablement de la même manière qu'il a conçu la solution, à partir du registre symbolique. Néanmoins il y aurait une autre voie de résolution, par le registre numérique. En effet, l'explication tentée par l'élève consiste à repérer dans le problème l'objet composé (A∩B, A∪B), puis la formule associée pour ensuite remplacer les expressions génériques par les données du problème. Cette démarche requiert :

- la connaissance de la formule,
- la connaissance du genre de situations auquel elle donne réponse,
- le repérage dans le problème des indicateurs du modèle (A∪B)

- une correspondance correcte (application) entre les données du problème et les paramètres de la formule

L'autre solution (Généralisation), consiste à se servir de la stratégie directe pour trouver les cardinaux, puis découvrir les régularités pour ensuite les généraliser, en arrivant à l'expression du registre symbolique.

Schémas de possibles solutions au montage B. Groupe 4

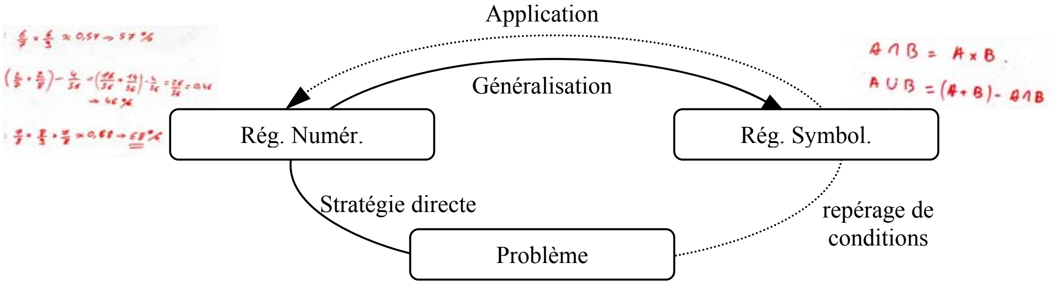


Figure 38

Les autres groupes ayant utilisé la stratégie indirecte pour les montages en série (A et C) ont suivi le schéma de l'application, mais d'une manière tacite. Le registre symbolique, en tant qu'algorithme est resté implicite pour ces élèves, ils n'ont pas eu besoin d'expliciter les étapes de calcul de ces probabilités. De cette manière, l'application représente la stratégie indirecte dont la spécificité consiste en la non explicitation de l'algorithme.

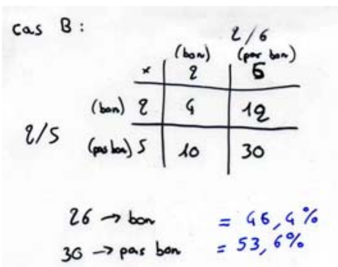
Le Groupe 4 semble avoir utilisé le sens de l'application. L'élève aurait identifié l'objet composé (A∪B) puis rappelé la formule respective pour finalement y substituer les valeurs des probabilités des composants. L'autre voie pour ce groupe nous semble moins probable :

- premièrement parce que l'élève essaye d'abord de justifier la formule ; une fois la formule admisse, son application devient immédiate,
- deuxièmement parce qu'il n'y a pas d'indices dans les feuilles de recherches que l'élève ait réalisé la stratégie directe, ce qui lui aurait permis d'expliquer le sens de chaque terme et
- troisièmement, parce que la formule ne fut pas demandée, si l'élève avait entrepris la voie de la généralisation, il se serait probablement arrêté à l'étape du registre numérique.

Enfin, la réponse de l'élève du Groupe 4 ne satisfait pas les élèves en général (min 74, « **Mais ça sert à rien cette formule** ») et la méthode est pratiquement écartée. Le Groupe 5 suit le Groupe 4. Ces élèves ont modifié à plusieurs reprises sa proposition, ils ont commencé par tenter une approche par la stratégie directe, postérieurement et en observant d'autres groupes ils ont penché pour l'indirecte, probablement pour la trouver plus proche des attentes du contrat didactique. Leur présentation est confuse, ils ont changé la méthode au fur et à mesure que les autres groupes présentaient les leurs, en passant de l'addition à la multiplication des proportions.

Finalement, et avant l'étape d'institutionnalisation, le Groupe 1 répond à la question que l'enseignant lui avait posé lors de sa présentation, il s'agissait de savoir comment ce groupe avait réussi à calculer les probabilités (stratégie directe). Le représentant de ce groupe montre un tableau à deux entrées pour le montage A (série), des croix symbolisant les pièces defectueses, des cercles pour les pièces non abîmées. La classe semble comprendre sans difficulté la stratégie de ce groupe. D'ailleurs, un autre élève du même groupe élabore la réponse pour les deux autres montages, toujours par la stratégie directe, cette fois-ci non pas en représentant les composant mais en indiquant les cardinaux respectifs. Ce transparent, dont on visualise les critères de comptage en fonction du type de connexion, n'est pas finalement présenté, et reste au sein du groupe.

Tableaux Groupe 1 montages B et C



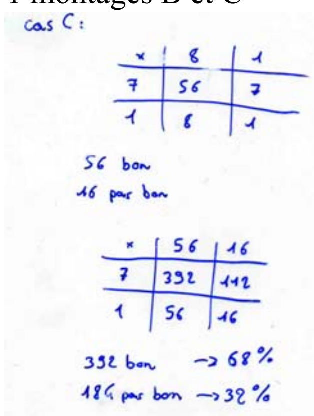
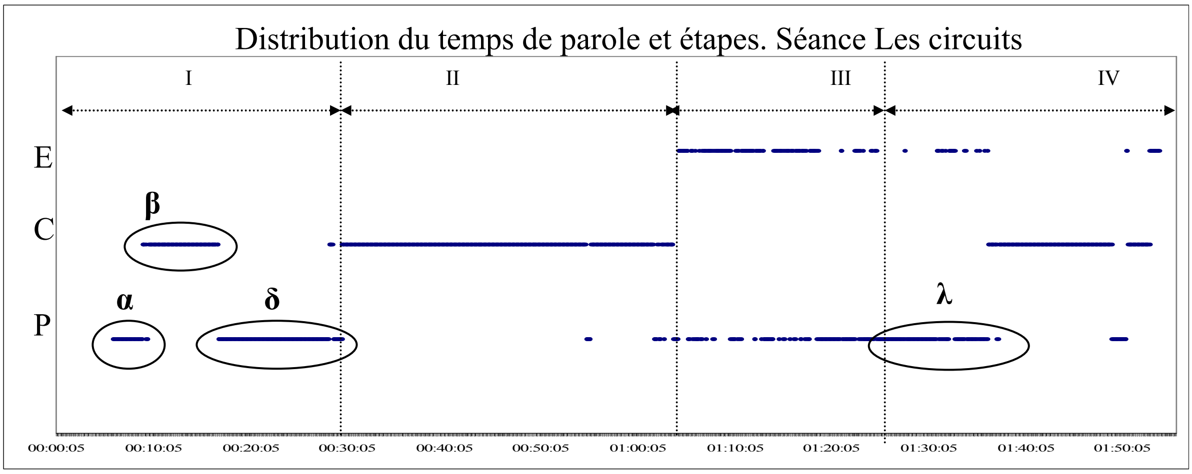


Figure 39

Institutionnalisation

Après la dernière présentation où le Groupe 1 montre sa méthode de calcul des probabilités, commence l'étape d'institutionnalisation (IV, voir Graphique). A partir de la minute quatre-vingts, et à la différence des étapes précédentes, l'enseignant guide le développement des idées. Dans la première partie de cette étape (λ, voir Graphique), l'enseignant vise la relation entre le signifié de la probabilité et l'évaluation par référenciation, les formules d'événements composés sont présentées comme un moyen efficace de calculer la probabilité par référenciation.



Graphique 46

La Figure 40 représente le plan envisagé par l'enseignant pour cette première partie de l'institutionnalisation. Il présente le terme probabilité associé au signifié de degré de certitude, en particulier pour ce problème, il le caractérise comme une mesure de certitude du fonctionnement des montages. Pour ce qui concerne la valeur numérique de cette mesure, il présente le rapport des cas favorables et possibles (principe d'équipossibilité) comme une raison à croire. Finalement, les formules d'événements composés comme un moyen économique de déterminer les valeurs nécessaires à ce raisonnement par référenciation.

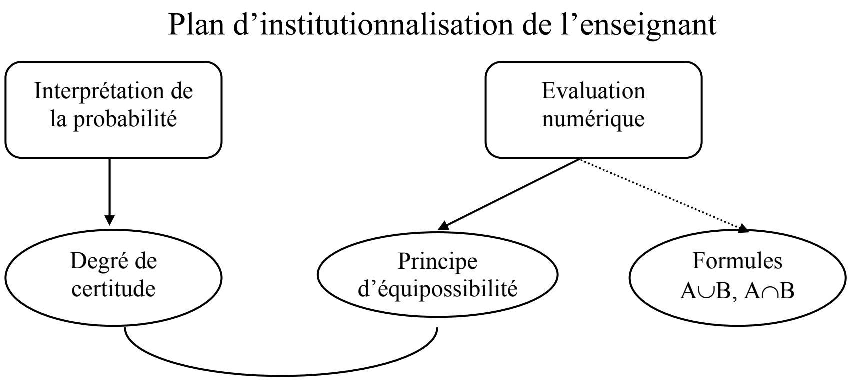


Figure 40

La classe ne parvient pas à traiter la relation entre les formules d'événements composés et la référenciation par des ensembles finis (lignée pointillée de la Figure 40). Le temps de cette institutionnalisation (30 minutes) est consacré à préciser quelques concepts et à analyser en détail les méthodes de détermination des cas favorables et possibles (tableaux, arbres, etc.).

Dimension sémantique

Dans l'étape d'institutionnalisation de cette séance, vu les difficultés des élèves à évaluer les probabilités, l'enseignant se voit obligé de simplifier les montages pour traiter les notions de base de l'articulation entre les aspects sémantiques et les calculatoires de la probabilité. Cette simplification consiste à réduire les montages à un seul composant. C'est la première tâche que l'enseignant effectue dans cette étape. En déconnectant les trois montages, il remet les pièces dans les pochettes respectives. Ensuite, et après avoir mêlé les composants tire une pièce de chaque enveloppe. Devant les sept composants sur la table, il repose la question du problème (min 84 : «*je voudrais* (…) *que mon ampoule s'allume vous me conseilleriez de prendre lequel parmi ces sept que j'ai sortie* »). Cette simplification lui permettra d'institutionnaliser tant la dimension sémantique de la probabilité (degré de certitude) que le critère de son évaluation par des ensembles finis de référence.

La question posée par l'enseignant réveille le débat sur l'approche encadrant la réponse. Un élève semble se situer dans le déterministe (Min 84 : « **Mais on ne peut pas savoir** »), un autre, dans l'indéterminisme (Min 85 : « **Bah sinon celui de plus haute probabilité** »). En fait, il semble que ce n'est pas la question proprement dite qui a réveillé ce débat (même si très court) mais plutôt le moment dans la séance dans lequel elle a été posée. En effet, cette même question a été posée au début de la séance (Min 23 : « *lequel de ces trois circuits vous me conseilleriez d'utiliser* »). Néanmoins, aucun groupe n'a discuté sur les valeurs logiques des objets concernés comme c'est le cas dans ce deuxième moment. A ce moment-là, les élèves se sont consacrés tout de suite à chercher des stratégies pour évaluer numériquement la probabilité.

Il semble que ces deux moments (min 23 et min 85) où l'enseignant pose la même question sont des moments bien différents. La tâche dévolue (Brousseau, 1998) aux élèves à la minute vingt trois ne serait pas la même que celle de la minute quatre-vingt-cinq, la première serait pour effectuer des calculs, la deuxième, pour s'exprimer. Les indices permettant de les différencier se trouveraient dans le contrat didactique construit en classe de mathématiques.

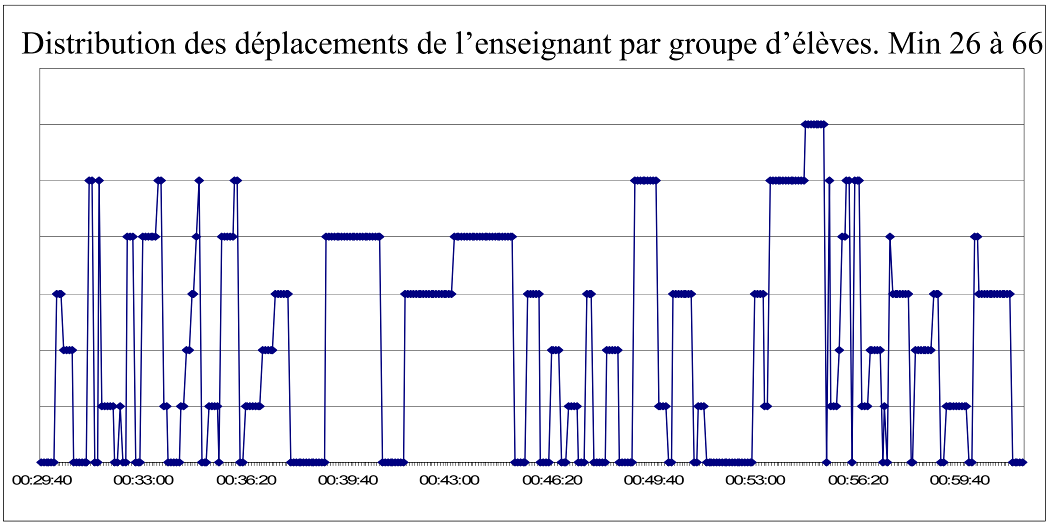
Il semble que la première de ces deux dévolutions (minute 23), où les élèves travaillent de manière autonome, serait perçue comme étant plutôt d'application de techniques, tandis que la deuxième, sous la tutelle de l'enseignant, serait un espace où les élèves pourraient s'exprimer sur des aspects épistémiques pour l'enseignant soit qui prend en charge la responsabilité de trancher les débats. De cette manière, il existerait dans ce deuxième moment une contre-dévolution dès élèves vers l'enseignant, l'enseignant rend une question aux élèves qui se permettraient d'y répondre si l'enseignant de sa part accepte la responsabilité de la gestion des débats et de l'élaboration des conclusions. Cette responsabilité de l'enseignant serait assurée par les caractéristiques mêmes de l'institutionnalisation.

Mais le débat concernant l'approche à choisir (déterminisme-indéterminisme) ne se prolonge pas longtemps, la classe semble se trouver déjà trop engagée dans les calculs des probabilités comme pour revenir sur ce genre de questions. Toutefois, l'enseignant profite des deux positions différentes pour introduire la probabilité bayésienne comme un indicateur de certitude partielle, un indicateur qui permet de prendre une décision dans des contextes d'incertitude.

La formalisation de ces notions reste su le plan du discours, les élèves ne sont pas invités à prendre de notes dans leurs cahiers (min 87 : « *On va bien formaliser au moins oralement* »). Le fait de ne pas laisser des traces de ces questions dans les cahiers des élèves pourrait s'expliquer par une sorte d'équilibre trouvé par l'enseignant. D'une part, il répond à nos attentes de traitement de ces notions, et d'autre part, vu qu'elles ne font pas partie du programme, il préfère ne pas charger les élèves en laissant des traces sur les cahiers avec des notions que ne serions pas évaluées. De cette manière, en faisant une institutionnalisation orale des aspects interprétatifs et décisionnels, il répondrait à nos besoins, sans trop altérer le programme de la classe.

Une autre caractéristique de cette étape où l'enseignant formalise certains concepts concerne les sources que nourrissent sa présentation. Son explication se réalise en général, à partir de réflexions propres basées sur des caractéristiques de la situation-problème. Par exemple, pour présenter la probabilité comme étant bayésienne, l'enseignant décrit l'épreuve générique en signalant qu'on s'intéresse à une caractéristique inconnue (fonctionnement ou non) d'un objet en particulier (un montage présent dans la salle). Pour expliquer le procédé de référenciation par des ensembles finis, il souligne le manque de connaissance sur l'objet en particulier (min 88) et l'appartenance de cet objet à un ensemble plus vaste dont on connaît la proportion d'objets ayant la caractéristique en question, etc. De cette manière sa présentation est la réunion d'un ensemble d'idées fruit de sa réflexion et des éléments caractéristiques du problème.

Dans cette présentation nous n'avons pas constaté que l'enseignant se nourrisse des productions des élèves provenant des étapes précédentes. Dans ce sens, il semblerait, vu la nouveauté du sujet, que l'enseignant souhaite s'assurer de présenter ces concepts le plus soigneusement possible, et que cette présentation a été préparée à l'avance et revisée durant la séance. Ainsi, devant ce sujet nouveau, l'enseignant aurait préféré s'appuyer principalement sur sa préparation pour institutionnaliser ces objets, plutôt que le faire à partir des idées des élèves, ce qui demanderait un travail additionnel de motivation, gestion et encadrement.



Graphique 47

Dans un passage de cette étape d'institutionnalisation, l'enseignant tente de faire participer la classe. Lorsqu'il est en train de commencer à présenter l'interprétation de la probabilité, un élève avance la réponse (min 85 : « **On a la probabilité que la pochette** »). Dans ce moment, l'enseignant essaie de reproduire un exercice que nous avons pratiqué lors de nos entretiens de préparation de la séance. Il s'agissait d'utiliser des termes synonymes du mot probabilité. Cet exercice visait l'explicitation par des termes synonymes du signifié que chacun de nous attribuions à la probabilité dans un moment donné. L'enseignant tente donc de le reproduire en classe :

86 *On essaie de tout dire sans prononcer le mot probabilité d'accord parce que l'idée qu'on a c'est*

d'essayer de se dire qu'est-ce que c'est cette probabilité donc on va essayer de le dire sans prononcer le mot probabilité
On va prendre la boite pour la quelle on a plus de chances de cas
Alors pas de probabilité pas de chances alors vas-y
Le maximum de cas favorables sur le nombre de cas total

Cette nouvelle consigne de l'enseignant n'est pas tout à fait comprise, l'élève en question complète sa réponse en donnant le critère décisionnel. Très probablement l'enseignant cherchait à reproduire notre exercice d'expressions synonymes. Mais il y a une différence fondamentale entre notre exercice et celui proposé à l'élève. Nous disposions, fruit de nos échanges, d'un certain nombre d'expressions synonymiques pour se substituer au terme probabilité, tandis que cet élève, pour répondre en se servant des expressions synonymiques devrait les avoir créés, cette tâche étant beaucoup plus complexe que la nôtre.

En résumé, l'enseignant continue avec la formalisation de l'interprétation de la probabilité et son évaluation numérique. Dans les minutes qui suivent la réponse de l'élève, l'enseignant retient plusieurs aspects liés à la probabilité. En décrivant le contexte comme étant d'incertitude, il considère la probabilité comme un outil décisionnel. Il rappelle, en d'autres termes, la nature de l'objet en question (min 87 : « *un objet il est là il est devant moi il faut que je choisisse parmi ces objets* ») et le critère d'évaluation de la probabilité (min 88 : « *je ne sais rien directement sur cet objet (...) par contre quand je prends un ensemble plus vaste un ensemble auquel appartient cet objet(...) je connais la proportion d'objets qu'ont la caractéristique que je voudrais que mon objet ait »).*

En suite, il demande aux élèves de relier les dimension sémantique et numérique pour deux probabilités : P(A)=0 et P(B)=1. Dans les deux cas le sens de la question est dès le registre symbolique vers le langagier (« min 88 : *qu'est-ce que ça signifie que la probabilité est égale à zéro* (…) »). Les réponses des élèves sont toujours en termes de "chances" : « **Je n'ai aucune chance** » et « **j'ai toutes les chances** » respectivement. C'est l'enseignant qui les reformule en termes de certitude (min 90 : « *Je n'ai aucune chance ça signifie que je suis certain* »).

Dans ce sens, les possibilités d'expressions alternatives pour le signifié de la probabilité sont assez restreintes chez les élèves. En général, un seul terme est utilisé comme synonyme de la probabilité, le mot chance, et cela tant pour la notion fréquentiste que pour la bayésienne. Cet espace réduit de termes utilisés par les élèves pour se référer à la probabilité est une difficulté à notre intention de repérage du sens attribué à la probabilité.

Néanmoins, les élèves semblent ne pas douter du signifié de ces mots lorsqu'ils discutent entre eux. Très probablement le contexte joue un rôle fondamental dans la communication entre les élèves. Ainsi, les ambiguïtés propres au terme probabilité seraient enlevées par l'information fournie par le problème. Nous reviendrons dans les conclusions sur ce sujet.

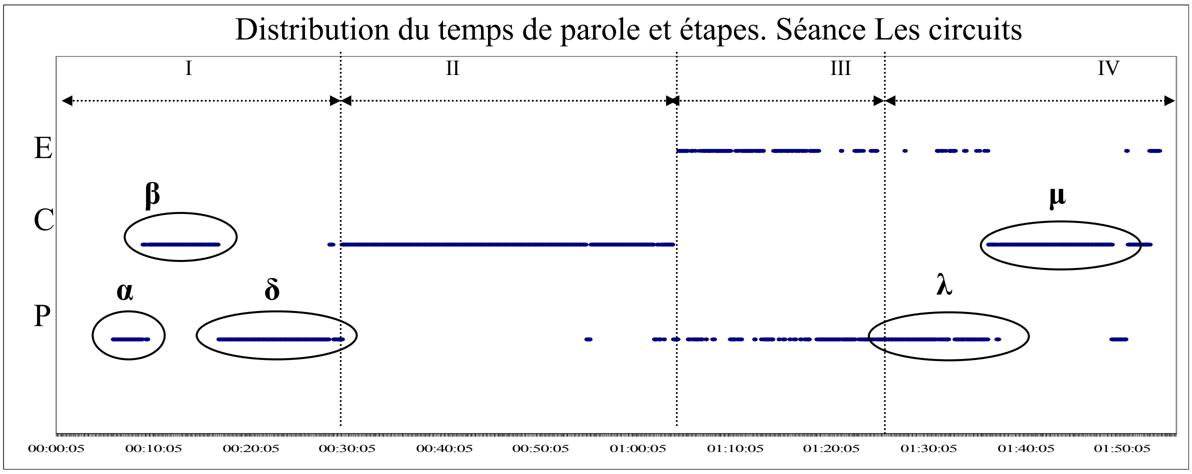
L'enseignant continue l'institutionnalisation, toujours autour de trois axes, a) signifié de la probabilité, b) la probabilité comme outil de prise de décision et progressivement c) son évaluation numérique. Ce dernier axe occupera le reste de la séance, c'est là où l'enseignant a constaté des difficultés et, vu son importance dans le programme, qu'il centrera toute l'attention du temps restant. Il explique que pour prendre une décision, on utilise la probabilité comme un indicateur et que cet indicateur s'évalue par (min 91) : «[la] *proportion de l'ensemble plus grand si je connais la proportion qui ont la caractéristique c'est en raisonnant sur cet ensemble plus grand que je vais que je vais calculer que je vais déterminer ce degré de certitude ».*

En suite il demande aux élèves s'ils ont raisonné de la manière par lui indiquée. Cette question, de l'ordre du métacognitif est répondue timidement par quelques élèves (« **un peu quand même** »). A cette question l'enseignant ajoute une condition qui place encore une fois les élèves en position métacognitive (min 93) « *non non on parle pas de probabilité* ». La question posée et la contrainte introduite (« *on parle pas de probabilités* ») demandent de réfléchir d'une part sur la méthodologie employée et d'autre part sur le signifié porté par le mot probabilité. Tâches qui, ensemble, deviennent insurmontables pour les élèves. L'enseignant donc revient sur le plan cognitif et demande de vérifier les calculs des probabilités de fonctionnement des montages en énumérant tous les cas possibles et favorables, c'est le début de la partie consacrée aux calculs (μ) de cette étape d'institutionnalisation.

.

Dimension calculatoire

A la minute quatre vingt sept les élèves commencent le travail en groupe (μ, Graphique 48) pour répondre à la question de l'enseignant. En général les groupes se servent de tableaux pour énumérer les cas favorables et défavorables. Les tableaux à deux dimensions leurs permettent de compter les cardinaux pour les circuits de deux composants, c'est donc sur les deux premier montages que les élèves effectuent l'énumération.



Graphique 48

Quelques groupes rencontrent des difficultés à intégrer les proportions dans leurs modèles. Par exemple, les Groupes 2 et 5 commencent en considérant quatre possibilités pour le montage en parallèle (B), dont trois cas favorable (Graphique 39), du même pour le montage en série dont la seule différence se trouve dans les cas favorables (un seul). Après l'intervention de l'enseignant, ces deux groupes trouvent la stratégie correcte en pondérant les possibilités par les proportions des pièces renfermées dans les pochettes.

.

Modèle d'énumération par possibilités

Figure 41

Ceci requiert un nouveau moyen de représentation dont les possibilités sont remplacées par les pièces. Ce dernier moyen de représentation est finalement compris par ces deux groupes, tous finalement parviennent à bien calculer les probabilités pour les montages à deux composants (A et B). Néanmoins, un seul (Groupe 1) réussi à adapter le tableau bidimensionnel au montage en série de trois pièces. Le Groupe 1 avait déjà résolu ce cas, ils avaient procédé à décomposer le problème en deux étapes, la première pour les deux premiers composants, puis ils avaient combiné ce résultat partiel avec le dernier composant. La technique de décomposition utilisée par le Groupe 1 est envisageable dans la pratique si au lieu de représenter les composants « bons » et « non bons » par colonne et ligne respectivement, on considère leurs cardinaux.

A la fin de la séance, ce groupe est invité à expliquer à toute la classe cette technique de décomposition en deux étapes. La séance se termine précisément avec cette présentation, à la minute cent quatorze. Nous présentons nos conclusions pour cette séance.

Conclusions

Nous avons conçu cette situation-problème pour explorer les possibilités de sensibilisation de la notion bayésienne de la probabilité. En particulier nous nous sommes intéressé à sa manifestation (ou émergence) en classe et sa institutionnalisation postérieure. Les premières conclusions de cette séance appelée *Les circuits* concernent les possibilités de son émergence, tirées principalement à partir du travail effectué en groupe par les élèves.

Sur l'interprétation bayésienne

Les indices provenant des échanges enregistrés nous amènent à penser que les élèves se sont bien représentés une épreuve générique. Les élèves se sont toujours référés aux objets présents dans la salle, les montages A, B et C. Leurs réflexions concernant le fonctionnement de ces montages étaient bien de caractère épistémique (« **c'est plus probable que çà marche** », « **que celui-là çà marche pas** »).Les élèves ayant évalué leurs probabilités par référenciation, nous n'avons trouvé aucune trace de la notion de fréquence d'apparition, ni en tant qu'objet principal ni comme raison de croire.

Ce repérage des éléments caractéristiques typiques de l'épreuve générique dans le discours des élèves nous a permis, de manière indirecte, d'identifier la notion de probabilité sous-jacente au discours des élèves. Néanmoins, toute tentative de recherche d'indices explicites ou directs par des expressions synonymiques nous a conduit à un échec. En effet, l'analyse du discours des élèves nous montre une diversité très limitée de termes alternatifs du mot probabilité. La principale option reste toujours le vocable chance.

Cette absence de mots alternatifs pourrait s'expliquer par une sorte de suffisance de ces termes. Autrement dit, dans cette séance, il n'y aurait pas eu besoin de recourir à d'autres mots que les deux indiqués (probabilité et chance). Cette suffisance serait atteinte par la contribution du contexte qui jouerait un rôle de complément du langage. Les ambiguïtés propres aux termes probabilité et chance seraient de cette manière enlevées par le contexte du problème. Bien que le signifié de ces termes soit dual, le sens ne serait que bayésien. Par exemple, le mot chance, ambigu dans son signifié, aurait un sens spécifique par les caractéristiques du contexte, prenant le sens d'"être plus sûr"ou "moins sur" de quelque chose.

D'ailleurs, lors de l'institutionnalisation, l'enseignant a essayé, en vain certes, de forcer l'emploi explicite de termes alternatifs (« pas de probabilité, pas de chance »). L'impossibilité des élèves à s'exprimer autrement que par ces deux termes même dans cette occasion, pourrait s'expliquer par la concurrence de deux facteurs, d'une part la difficulté à se situer en une position métacognitive, d'autre part le procédé d'introspection propre à la notion bayésienne.

- Pour ce qui concerne les métaconnaissances. La recherche de termes substituts à la probabilité bayésienne requiert, de la part des élèves, une position réflexive sur le sens attribués par eux-mêmes au mot probabilité. La complexité de cette tâche pourrait expliquer les difficultés rencontrées pour trouver des indices explicites de cette notion chez les élèves (Robert & Robinet, 1993).
- Pour ce qui concerne l'introspection. La notion bayésienne de la probabilité, en tant que degré de certitude, est une appréciation personnelle sur la véracité d'une proposition donnée. La mesure de cette appréciation décrit une caractéristique de la personne et non pas de la proposition. Pour assigner une valeur, l'individu doit observer une caractéristique propre à lui, dans ce sens nous parlons d'introspection. Bien que ce processus d'évaluation se voit facilité par la naturalité de la référenciation par des ensembles finis, la conceptualisation qui le rend conscient semble difficile aux les élèves.

  Pour la notion fréquentiste par contre, les termes synonymes sont apparus plus facilement lors de la première séance (rare, fréquent, etc.). L'extrospection (notion fréquentiste) serait donc plus accessible aux élèves que l'introspection (notion bayésienne).

De cette manière, plusieurs facteurs semblent concourir pour que la recherche d'expressions alternatives soit finalement difficile pour les élèves. Le contexte donc semblerait avoir une place particulièrement importante dans la communication d'idées, dans ce genre de circonstances dont le lexique est restreint, au même temps qu'ambigu.

Sur l'indéterminisme

Les débats concernant les possibles paradigmes métaphysiques (déterminisme, indéterminisme) ont été moins présents dans cette séance que dans la première. Une explication pourrait se trouver en ce que, tant le paradigme indéterministe que le critère décisionnel, auraient été naturellement acceptés par les élèves. De cette manière, il n'y aurait pas eu motif à discussion. Néanmoins, une autre explication nous semble possible, celle de type de tâches habituellement à la charge des élèves, un type lorsqu'ils travaillent de manière autonome (produits) et un autre lorsqu'ils le font sous le guide de l'enseignant (processus). Dans cette séance, il n'y a qu'un seul échange (entre deux élèves) concernant la position à prendre devant l'incertitude. Ce dialogue, très court d'ailleurs, s'est produit dans un moment où l'enseignant non seulement coordonnait les discussions mais aussi les guidait.

D'ailleurs, en observant les conditions dans lesquelles se sont produits les débats lors de la première séance nous constatons la même caractéristique : l'enseignant modère et guide les propositions des élèves. En d'autres termes, toutes les discussions concernant les critères décisionnels se sont déroulé sous la tutelle de l'enseignant. Les difficultés en se situer en position métacognitive et une distribution de tâches par le contrat didactique pourraient expliquer ce phénomène.

Les élèves auraient donc tendance à travailler de manière autonome sur les produits, et de manière guidée sur les processus. Les processus étant des démarches métacognitives, les produits, cognitifs. Ainsi, les élèves éprouveraient plus des difficultés dans les premiers (arguments décisionnels, introspections, etc.) que dans les deuxièmes (calculs des probabilités).

Sur l'institutionnalisation

Il nous intéressait aussi pour cette séance d'analyser les possibilités d'institutionnalisation de la notion bayésienne de la probabilité, Bien entendu, dans un contexte où cette notion n'a pas une reconnaissance formelle. Un effet observable de cette non reconnaissance peut se trouver dans le registre dans lequel la probabilité bayésienne est formalisée. Dans cette séance, l'enseignant choisi le langage oral pour toutes ses synthèses concernant l'interprétation de la probabilité et les prises de décision. Ce choix de l'enseignant nous semble le résultat de sa nécessité de répondre à deux genres de demandes. D'une part nous et notre intérêt pour analyser les possibilités d'attribuer aux deux notions de la probabilité le même statut institutionnel, de l'autre part, le programme officiel qui le contraint à en enseigner (et évaluer) qu'une. Ainsi, par une institutionnalisation dans ce registre, l'enseignant répondrait, au moins partiellement, aux deux demandes : nous pourrions étudier les possibilités de traitement de cette notion, et les élèves ne garderaient des traces que de ses aspects qui concernent le programme.

Dans cette séance, donc, l'institutionnalisation des aspects sémantiques de la probabilité se réalise dans le registre langagier, sans laisser des traces dans les cahiers des élèves.

Une autre caractéristique de l'institutionnalisation des aspects sémantiques concerne les sources utilisées (au moins explicitement) par l'enseignant. Il se nourrit principalement de ses propres réflexions et des caractéristiques du problème. Sa présentation détaillant tant les aspects sémantiques que calculatoires ; les termes utilisés (y compris des synonymes) et le temps consacré (cinq minutes), nous parlent d'une présentation soigneusement préparée : l'enseignant présente la probabilité en tant que degré de certitude, la référenciation par des ensembles finis, et même il n'oublie pas de mentionner le critère de maximisation de la probabilité comme critère décisionnel. Néanmoins, il ne reprend pas les échanges des élèves lorsqu'ils ont travaillé en groupe.

Une explication à ce phénomène pourrait se trouver dans la présence presque nulle de cette problématique dans ces échanges. En effet, en observant les enregistrements de cette étape nous constatons une centration sur des aspects calculatoires.

Sa préoccupation pour une présentation soignée, la nouveauté du sujet et notre présence en classe semblent mieux expliquer la non prise en compte des débats des élèves.

Pour ce qui concerne les aspects calculatoires, leur institutionnalisation est restée inachevée, de la même manière que lors de la première séance. Dans la deuxième partie de cette dernière étape de la séance, l'enseignant renvoie aux élèves la tâche de détermination des cardinaux des ensembles nécessaires au calcul de la probabilité. Néanmoins, ce travail des élèves ne parvient pas à être mis en commun par l'enseignant qui partage une partie de cette tâche avec Groupe 1, seul à avoir trouvé un moyen pour calculer la probabilité du montage en parallèle, probablement parce qu'en visitant les groupes il perçoit la méthode déjà comprise.

Pour finir l'analyse *a posteriori* de cette séance nous transcrivons une partie des impressions générales que l'enseignant nous a fait parvenir à la fin de la séance : (min 106, fin de la séance) :

> *« (...) ce qui m'a un peu fasciné dans les débats c'est qu'il y a des groupes il y avait des groupes qui avaient dénombré quoi et ça pouvait servir pour d'autres et dire de le faire et dire tiens est-ce que finalement parmi ces calculs que je faisais qui apparaissent qui va me permettre de lui donner un peu de sens quoi de lui donner du sens Jeremy je trouve ça excellent ce qu'il a fait c'est vachement bien ça (...) je pensais à (...) on passe un temps (...) à la formule (...) et on n'a rien (...)donc je me dis ça vaut peut-être le coup j'étais plutôt satisfait parce que j'ai vu qu'il n'y a eu aucun qui n'est pas rentre*

dedans (...) quoi moi j'avais peur qu'ils me disent attendez on a fait des calculs quoi c'est bons les calculs on les a faits on les connaît les calculs mais non le fait de le faire pour trouver quelque chose ça les à conforté (...) parce qu'ils ont vu des choses de probabilité et là ils ont des décisions à prendre alors que jamais on leur demande de prendre une décision et on leur dit pas comment faire ou on leur laisse libre (...) ce qui m'a un peu fasciné dans les débats c'est que (...) il y avait des groupes qu'avaient dénombré (...) ce qui va me permettre de lui donner un peu de sens (...) ».

Par la suite nous présentons la troisième et dernière expérimentation. De cette manière nous complétons les trois types de nature d'objets ; la première a été sur la notion fréquentiste de la probabilité, l'objet étant la série infinie, la deuxième pour la notion bayésienne, où l'objet est une épreuve générique ; la troisième finalement sera aussi pour la notion bayésienne où l'objet est une hypothèse. Dans la prochaine section nous présentons cette troisième expérimentation.

4.1 Analyse de troisième expérimentation

Cette troisième situation-problème correspond au type d'objet *hypothèse* de la notion bayésienne de la probabilité, sa préparation, comme les autres, a suivi plus ou moins celle de cette synthèse : d'abord la détermination des caractéristiques générales, puis la recherche d'un problème en accord avec les caractéristiques en question. Réunir toutes ces conditions dans un problème n'a pas été facile à réaliser. Après plusieurs tentatives, nous avons trouvé à la situation proposée par Guy Brousseau (Brousseau et al., 2001) un double intérêt, non seulement cette situation pouvait nous servir de de base pour expérimenter la notion bayésienne dans sa modalité hypothèse, mais aussi, en la modifiant, elle nous permettait de discuter quelques différences entre les deux notions de la probabilité.

Au début nous avions séparé ce problème en deux parties, une bayésienne et l'autre fréquentiste. De cette manière, les deux notions pouvaient être traitées dans un "même problème". La deuxième partie (fréquentiste) différait de celle proposée par Guy Brousseau par l'objet sur lequel porte la probabilité. En effet, dans la situation de Guy Brousseau, l'objet est la série des couleurs des billes apparues lors des retournements de la bouteille, on s'intéresse à la fréquence d'apparition de chaque couleur lorsque l'on retourne la bouteille. Dans notre première partie, bayésienne, la probabilité porte sur la bouteille présente dans la salle, dans la deuxième partie la probabilité porte sur une série de bouteilles. De cette manière nous cherchions à mettre en évidence les différences entre les deux interprétations l'une sur une bouteille, l'autre sur un ensemble de bouteilles.

Finalement, et pour des raisons diverses nous avons renoncé à expérimenter la deuxième partie. La situation-problème que nous avons expérimenté est donc la première partie consacrée à la notion bayésienne de la probabilité, celle qui probabilise sur la bouteille.

La bouteille

Analyse a priori

Matériel

- Une bouteille opaque () avec un bouchon transparent. Un sac rempli de billes de deux couleurs.

Figure 42

- Des ordinateurs équipés avec un tableur.
- Vingt jetons ou pièces de monnaie par binôme.

- Une feuille contenant un tableau (**Tableau 58**) par binôme.

Bille	Pièces de monnaie par composition					Justifiez la distribution des pièces
	A NNNN	B NNNO	C NNOO	D NOOO	E OOOO	

Tableau 58

Le problème

L'enseignant distribue une feuille (**Tableau 58**) et vingt jetons à chaque binôme d'élèves. Puis il montre une urne remplie de billes de couleur noire et orange. La composition de cette urne étant inconnue aux élèves, il retire quatre billes de l'urne ; sans les montrer, il les introduit dans la bouteille. Par la suite, la seule action possible sera d'effectuer des retournements de la bouteille pour observer une bille à la fois à travers le bouchon transparent.

Le professeur propose d'analyser une méthode permettant d'estimer la composition de la bouteille sans l'ouvrir et à chaque fois qu'il retourne la bouteille. L'estimation s'exprimera en termes de nombre de jetons assignés à chaque composition possible. Pour cela, il demande de distribuer les vingt jetons selon les consignes suivantes (ces consignes se trouvent sur les feuilles des élèves) :

- Plus vous croyez en une composition, plus vous lui assignez de pièces.
- Toutefois, si vous croyez plus en une composition qu'en une autre, vous devrez le justifier.
- Lorsque vous êtes d'accord sur la distribution des pièces, vous écrivez dans les casiers respectifs le nombre de pièces que vous avez assigné à chaque composition.
- Un exemple : Si vous n'assignez aucune pièce de monnaie à la composition C, cela veut dire que vous êtes sûrs que la bouteille n'est pas du type NNOO. De la même manière, si vous assignez toutes les pièces à la composition D, c'est parce que vous êtes sûrs que la bouteille est du type NOOO.

Objectif pour l'expérimentation

Dans cette séance, nous cherchons à expérimenter sur la possibilité d'aborder en classe la notion bayésienne de la probabilité. En particulier nous souhaitons sensibiliser les élèves sur le type hypothèse et son évaluation numérique.

Solution

Par ses caractéristiques, ce problème correspond à un test d'hypothèse bayésien, et non pas fréquentiste. Voici quelques raisons :

- Le nombre d'hypothèses. Un test d'hypothèse fréquentiste est tout d'abord dichotomique (H_0 et H_1). Notre problème propose d'analyser cinq hypothèses de manière conjointe.
- Les étapes. A l'origine des tests fréquentistes se trouve l'intention d'étudier un effet de traitement sur une population donnée. De cette manière, et grâce au traitement, la population passerait d'une hypothèse H_0, supposée connue ou admise, à une autre H_1. Ce nouvel état serait attribuable au traitement effectué (Hacking, 1965; Hacking & Dufour, 2004; Neyman, 1977). En principe, les problèmes abordés par ce genre de tests s'intéresseraient donc à un possible changement dans les paramètres d'une population, elle passerait d'un état H_0 à, éventuellement, un nouvel état H_1. Pour les tests bayésiens la situation sera moins restrictive que celle de l'approche fréquentiste : on assigne des probabilités aux hypothèses et on apprend sur elles au fur et à mesure que l'information s'accroît.
- Dans ce sens il y a une différence épistémique entre le test fréquentiste et le test bayésien, différence qui aura des conséquences sur la préparation de situations-problèmes. Cette différence épistémique vient de ce que, pour un test fréquentiste et avant l'effet de traitement, le chercheur dispose d'une certaine connaissance sur la population lui permettant d'accepter l'hypothèse H_0. Une conséquence de cette relation épistémique sera qu'il faudra concevoir le problème de telle manière que les élèves puissent disposer d'information sur la population. Cette information devrait faire que l'hypothèse H_0 soit acceptée. Il nous semble que l'acceptation d'une hypothèse nulle sans être fondée affaiblirait à la fois le procédé argumentatif du test et la confiance qu'on porterait sur l'éventuelle nouvelle hypothèse.

  Dans notre situation-problème, les élèves ne disposent d'aucune information leur permettant ni d'accepter une hypothèse nulle ni d'en proposer une alternative particulière.

Le genre de rapport épistémique caractérise le genre de problèmes associé à chaque approche (fréquentiste ou bayésienne) au moins d'un point de vue didactique. En principe, l'état épistémique des hypothèses de ce problème mène à le modéliser d'un point de vue bayésien. Cinq compositions sont possibles pour la bouteille :

Possibles compositions de la bouteille

A	B	C	D	E
NNNN	NNNO	NNOO	NOOO	OOOO

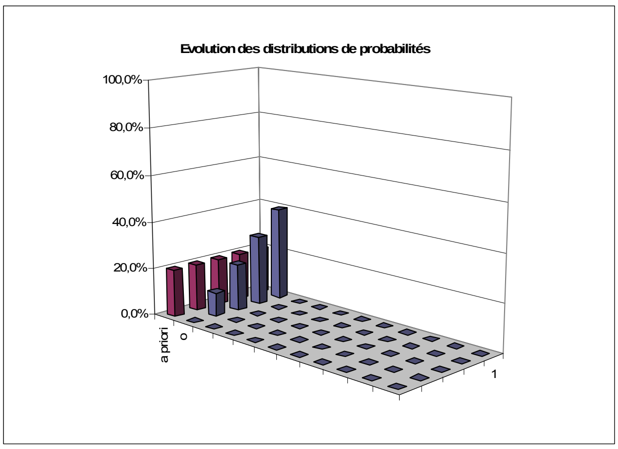
Tableau 59

N : bille noire. O bille orange

Le manque de connaissance sur la proportion de billes dans l'urne conduit à chercher d'autres raisons pour assigner les jetons, parmi ces raisons se trouve le principe de raison insuffisante. D'après ce principe, la même probabilité devrait être assignée à toutes les compositions. L'évolution des probabilités se fera en fonction des couleurs des billes observées lors des retournements de la bouteille.

Par exemple, si lors du premier retournement on observe une bille de couleur orange, on peut écarter l'hypothèse A (NNNN). Son élimination de l'ensemble des possibles se réalise par déduction, néanmoins, celle-ci n'est pas la seule conclusion possible. Une bille orange est un phénomène rare pour la composition B, et beaucoup plus fréquente pour la D, de cette manière il semblerait plus probable que la bouteille soit du type D que du type B. Ce genre de raisonnement est connu comme une abduction. De cette manière, on devrait croire plutôt aux hypothèses où ce phénomène est plus probable à observer qu'à celles où il est rare.

Le taux de changement pour chacune des probabilités est complexe à déterminer, c'est la formule de Bayes qui donne une réponse à cette question. En appliquant cette formule, nous obtenons la distribution de probabilités du **Graphique 49**.



Graphique 49

$$P_O(H_j)=\frac{P_{Hj}(O)}{P(O)}\times P(H_j) \quad (1)$$

où :

$P(H_j)$: la probabilité portée sur l'hypothèse H_j avant l'arrivée du fait O; avec j compris entre 1 et n.probabilité *a priori*)

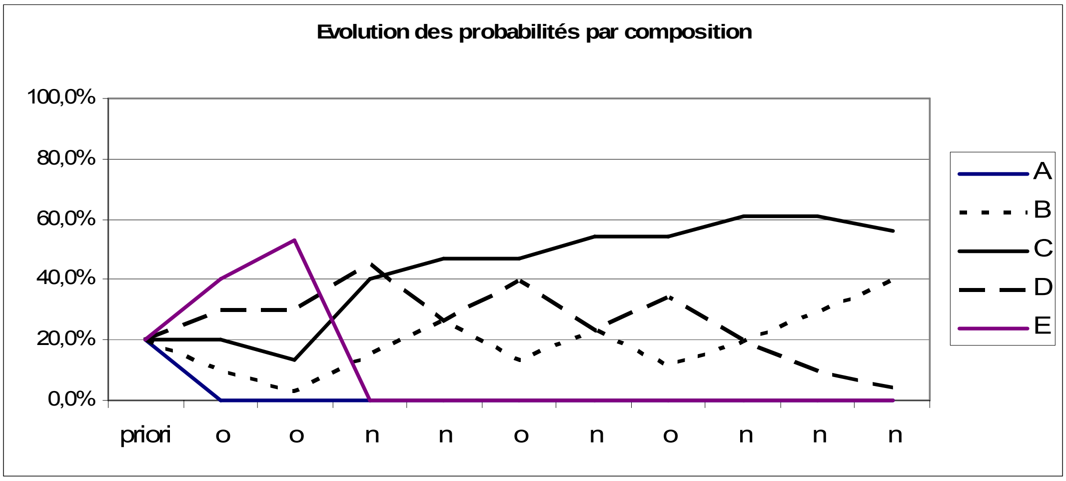
$P_{Hj}(O)$: la probabilité de O sous l'hypothèse H_j

$P(O)$: la probabilité globale (ou totale) d'arrivée du fait O ($\sum_{j=1}^{n} P_{Hj}(O)\times P(H_j)$)

$P_O(H_j)$: la probabilité portée sur l'hypothèse H_j après avoir appris O (probabilité *a posteriori*)

La formule de Bayes permet de compléter un cycle, on passe d'un état de probabilités *a priori* à un autre d'*a posteriori*. Il est possible, vue l'indépendance des retournements, de répéter ce cycle une autre fois, nous avons choisi d'effectuer dix retournements, ce qui renvoie à dix itérations de la formule de Bayes.

La solution au problème de la bouteille par le critère de probabilité maximale est relative aux couleurs obtenues lors des dix retournements. Le **Graphique 50** représente l'évolution des probabilités pour une série de quatre orange et six noirs (OONNONONNN). A la fin de ces dix retournements, et partant d'un état d'ignorance (équiprobabilité) la composition la plus probable est la C (OONN).



Graphique 50

Signalons que l'intérêt n'est pas seulement de déterminer la composition la plus probable au bout de dix lancers. Cela est relativement facile, il suffit de déterminer la proportion de couleurs obtenues dans les dix retournements et de la comparer avec les proportions des compositions possibles, de cette manière on choisira la composition dont la proportion est la plus proche à l'échantillon obtenu.

En fait l'intérêt de ce problème est de trouver un moyen pour réassigner les probabilités au fur et à mesure qu'on apprend sur la composition grâce aux retournements de la bouteille.

Déroulement

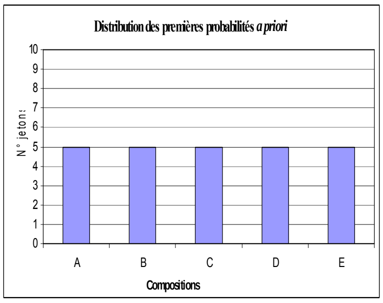
Première et deuxième étape : Introduction et représentation par jetons

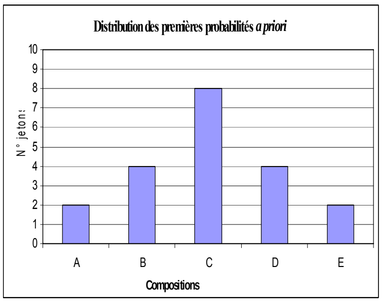
L'enseignant distribue le matériel pour la séance : une feuille contenant le **Tableau 58** et vingt jetons à chaque binôme d'élèves. Il montre l'urne de manière telle que les élèves puissent confirmer qu'à l'intérieur il n'y a que des billes orange et noire. Puis, il explique les consignes.

L'enseignant ne doit pas informer sur la proportion de billes dans l'urne. Ce choix, fonctionnant comme une variable didactique, empêche les élèves de se représenter cette première distribution comme une épreuve générique. En effet, si au début du problème, l'enseignant indique les proportions dans l'urne, les élèves y trouveraient un ensemble de référence fini pour évaluer leurs probabilités. De cette manière, les premières évaluations de probabilités s'effectueraient par référenciation.

Nous avons choisi pour cette occasion d'expérimenter le principe de raison insuffisante. L'enseignant refuse donc de donner toute information concernant la composition de l'urne. La classe doit trouver d'autres raisons que les proportions dans l'urne pour distribuer les jetons.

Plusieurs distributions sont possibles pour cette première étape. La **Figure 43** représente trois modèles, tous symétriques. Le premier serait justifié par le principe de raison insuffisante, le deuxième et le troisième par l'entropie (ou désordre) de chaque composition. Pour le deuxième modèle, les élèves analyseraient la séquence de couleurs de chaque composition et assigneraient plus de probabilité à celles qui leur semblent plus représentatives d'une phénomène aléatoire. Par exemple, en comparant les compositions C (NNOO) et B (NNNO), ils vont trouver la prmière plus désordonnée (plus grande entropie), donc plus probable.





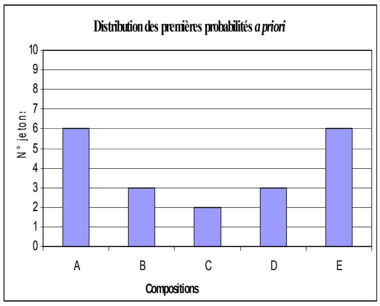
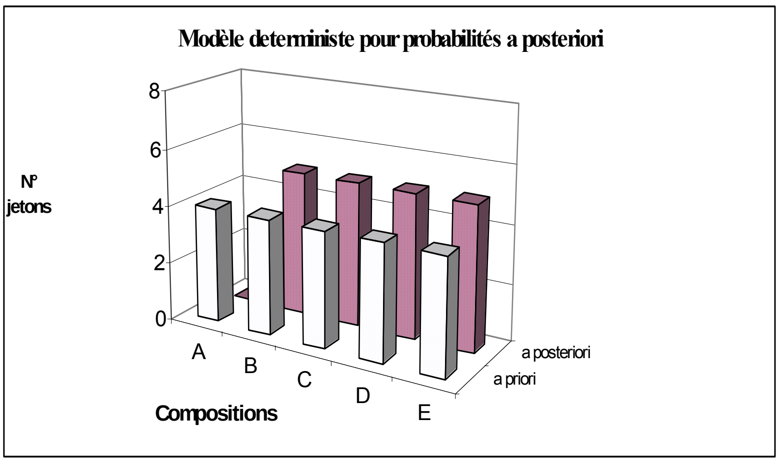


Figure 43

Une fois que les binômes d'élèves ont choisi et justifié leur distribution, l'enseignant retourne la bouteille. Pour n'importe quel des trois modèles, la couleur de la première bille implique l'élimination d'une des hypothèses monochromatiques (A : NNNN ou E : OOOO). Les élèves doivent redistribuer au moins les jetons enlevés de l'hypothèse impossible.

Plusieurs réactions sont possibles. Par exemple, admettons qu'un binôme ait choisi le premier modèle pour distribuer ses jetons et que la première bille observée ait été orange. L'hypothèse A étant impossible, ce binôme doit replacer les quatre jetons enlevés de A sur les autres compositions. Une possibilité consiste à assigner un jeton à chacune des autres hypothèses restantes. De cette manière ils auront une nouvelle distribution équiprobable, cette fois-ci pour les probabilités *a posteriori* (**Graphique 51**).

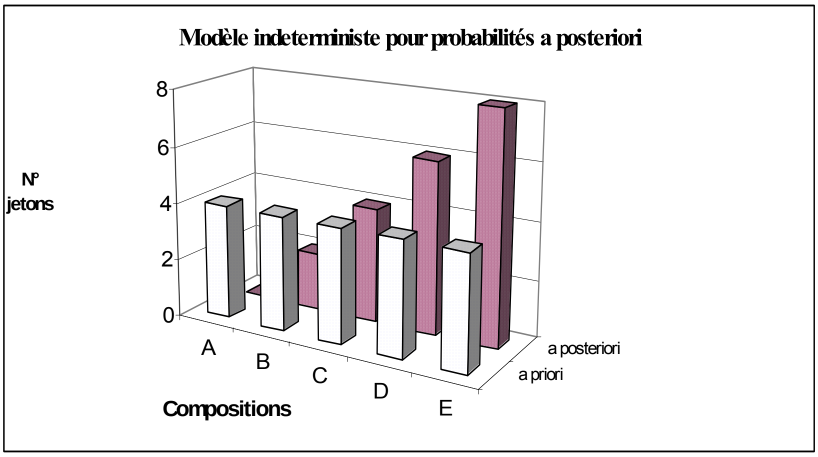
Néanmoins, la distribution équiprobable *a posteriori* aura un signifié bien différent celui de de la première. En effet, si les élèves choisissent cette distribution *a posteriori* cela voudra dire qu'ils n'abordent que la conclusion déterministe : l'hypothèse impossible. Il n'y aurait pas, dans ce cas, une mobilisation du raisonnement par abduction où une hypothèse devient plus probable qu'une autre par son caractère explicatif du phénomène observé.



Graphique 51

Une réponse par l'approche indéterministe (abduction) sera celle d'une distribution *a posteriori* du type du **Graphique 52** (**Tableau 60**) où non seulement on retire les quatre jetons de l'hypothèse devenue impossible mais aussi quelques uns (deux) de la composition B (NNNO). Cette dernière composition devient, après la bille orange, moins probable qu'auparavant.

La composition C (NNOO) reste de probabilité constante lors du premier retournement. Pour cette hypothèse, et vue l'état de connaissances sur le système (distribution *a priori*), une bille orange n'apporte rien de significatif. De leur côté, les compositions D et E verronnt augmentées leurs probabilités.



Graphique 52

	Compositions				
	A	**B**	**C**	**D**	**E**
Dist. a priori	*4*	*4*	*4*	*4*	*4*
Dist. a poster.	*0*	*2*	*4*	*6*	*8*

Tableau 60

La formule de Bayes permet de quantifier cette redistribution de jetons. En effet, en appliquant cette formule par exemple à la composition B nous obtenons :

$$P(O) =$$

$$P_A(O) \times P(A) + P_B(O) \times P(B) + P_C(O) \times P(C) + P_D(O) \times P(D) + P_E(O) \times P(E)$$

$$P(O) = \left(0 + \frac{1}{4} + \frac{1}{2} + \frac{3}{4} + 1\right) \times \frac{1}{5} = \frac{1}{2}$$

$$P_O(B) = \frac{P_B(O)}{P(O)} \times P(B) = \frac{\frac{1}{4}}{\frac{1}{2}} \times \frac{1}{5} = \frac{1}{10}$$

Nous ne cherchons pas, dans cette séance, que les élèves mobilisent des tels recours au programme de lycée. Nos objectifs sont :

- Faire émerger le raisonnement abductif qui fournit les arguments pour le déplacements des jetons et
- Créer la nécessité d'une méthode pour recalculer les probabilités. La formule de Bayes, introduite par l'enseignant, sera une réponse à cette nécessité.

La représentation de la probabilité sous la forme de jetons et les deux distributions sollicitées cherchent à faciliter la concrétisation de ces deux objectifs. Avec la représentation par jetons nous cherchons à libérer les élèves des tâches calculatoires pour qu'ils puissent se concentrer sur les principes régissant les assignations. Avec l'utilisation de jetons pour deux retournements nous cherchons à ce que le raisonnement abductif soit convoqué à plusieurs reprises.

D'ailleurs, lors du deuxième retournement, la précision de la représentation par jetons deviendra très probablement insuffisante, les élèves trouveronnt dans le registre numérique un système de représentation plus approprié. La nécessité de précision impliquera non seulement un moyen de représentation précis mais aussi une méthode pour déterminer les probabilités *a posteriori*, la formule de Bayes.

Troisième étape : Premier bilan et révision des distributions

Cette deuxième étape commence après avoir conclu la deuxième distribution *a posteriori*. Etant donné que les objectifs de mobilisation du raisonnement par abduction et de passage au nombre auront été moyennement atteints, l'enseignant devra intervenir pour introduire une étape de bilan. Dans cette étape, nous prévoyons une mise en commun de plusieurs notions :

La probabilité comme signifiant du degré de certitude.

Entre autres, cette situation-problème se différencie des deux précédentes par la manière d'introduire la notion de probabilité. Dans le problème du jeu des pièces de monnaie comme dans celui des circuits nous n'avions fait aucune référence à la probabilité. Dans ces deux problèmes, nous avions juste demandé des prises de décisions, la probabilité devait, en principe, être mobilisée par les élèves afin de trouver une réponse à ces demandes. La situation dans cette troisième expérimentation est différente. Dans la présentation du problème nous introduisons le signifié de la probabilité (« plus vous croyez… ») et aussi un moyen de représentation (les jetons).

Néanmoins, à ce concept manque son signifiant dans le registre langagier (le mot probabilité) et un registre performant (le numérique). Trouver son signifiant et sa représentation dans le registre numérique et symbolique sont des tâches que nous attendons de la part des élèves. Un objectif pour cette étape consiste donc à une mise en commun des signifiants utilisés dans la séance et du registre numérique comme moyen d'évaluation. Dans cette étape, le concept de probabilité bayésienne sera, en quelque sorte, complété, l'énoncé aura apporté le signifié et la classe le signifiant et un registre pour le quantifier.

Une fois la deuxième distribution *a posteriori* complétée, l'enseignant, réalise la mise en commun sur les plans signifiant et signifié de la probabilité bayésienne associée à ce problème.

La représentation numérique comme registre précis d'évaluation

Une autre mise en commun concerne le passage d'un espace de représentation discret à un autre continu, des jetons aux nombres réels. La nécessité d'un moyen de représentation plus performant que les jetons devrait se manifester lorsque les élèves essaient de déterminer la deuxième distribution *a posteriori*. En fait, nous nous attendons à que quelques binômes n'arrivent pas à établir cette distribution telle qu'ils la souhaitent. De cette manière, la contrainte se transforme en nécessité. La représentation numérique apparaît comme réponse à cette nécessité.

Soulignons que la représentation par jetons peut rester suffisante dans certains cas. Le **Tableau 61** montre que pour une distribution basée sur le principe de raison insuffisante et en réassignant les probabilités par la formule de Bayes, si les billes obtenues sont de couleurs différentes alors la distribution par jetons est suffisante avec deux retournements seulement. A fin de réduire le risque de l'occurrence de couleurs différentes, nous avons introduit dans l'urne soixante dix pour cent de billes orange et trente pour cent de noire.

	Pièces de monnaie par composition				
Bille	A NNNN	B NNNO	C NNOO	D NOOO	E OOOO
O	4	4	4	4	4
	0	2	4	6	8
N	0	6	8	6	0

Tableau 61

Toutefois, nous nous attendons que quelques binômes éprouvent des difficultés à représenter leurs probabilités par des jetons, au moins dans une des distributions *a posteriori*. Dans cette étape la classe devrait retenir le registre numérique comme nouveau moyen de représentation.

Le principe de raison insuffisante pour évaluer la première distribution *a priori*

Un autre concept à formaliser est celui du principe de raison insuffisante. Pour ce genre de propositions on doit y adhérer par le pois de ses arguments. En d'autres termes, ce principe n'étant pas déductible, sa validité es conséquence de l'acceptation par la cohérence de ses arguments.

La tâche de l'enseignant est ici délicate, par la place qu'il occupe dans la classe il pourrait imposer ce principe facilement, néanmoins, nous cherchons à ce qu'un minimum de débat soit établi en classe, un débat permettant aux élèves d'arriver à un consensus sur la cohérence des arguments de ce principe.

La classe devrait connaître ce critère non seulement pour définir la distribution associée (équi-répartie) mais aussi les conditions de son application (absence d'information autre que le nombre d'hypothèses possibles).

Les principes ou critères du raisonnement par abduction

Ayant formalisé la notion de probabilité, le registre de représentation et le critère pour la distribution *a priori*, l'étape suivante consiste à caractériser le raisonnent par abduction. Ce raisonnement permet le passage de la distribution *a priori* à celle *a posteriori*. L'argument pourrait se résumer comme étant la confirmation d'une hypothèse par son pouvoir d'explication d'un phénomène observé. Dans ce genre de raisonnement on avance des données vers les hypothèses. C'est en appliquant ce raisonnement qu'on peut apprendre sur la composition de la bouteille.

Cette étape de premier bilan a comme fonction, entre autres, celle de reconnaissance intitutionnelle de ce genre de démarches.

Introduction de la formule de Bayes pour calculer les probabilités *a posteriori*.

Ayant déterminé les deux distributions *a posteriori*, l'enseignant introduit la formule de Bayes comme un mécanisme de calcul qui automatise le procédé abductif. De cette manière la formule devient un outil pour déterminer avec précision les probabilités *a posteriori,* puis dans des futures séances elle pourra être analysée plus en détail.

Quatrième étape : Passage à l'ordinateur

Pour continuer à estimer la composition de la bouteille, l'enseignant propose d'effectuer des nouveaux tirages, jusqu'à dix au total. Les jetons devenant insuffisants et les calculs trop complexes, la classe continue le travail sur l'ordinateur. Pour cela, nous aurons enregistré sur chacun des ordinateurs un fichier tableur préparé à l'avance (**Figure 44**). Sur ce fichier les élèves indiquent leurs probabilités *a priori* (boutons jaunes de la feuille) et les couleurs obtenues (colonne à gauche de l'écran). La feuille, étant protégée, calcule les probabilités *a posteriori* pour chaque composition possible en fonction de la distribution *a priori* de la couleur de la bille visualisée. Les graphiques permettent de suivre l'évolution des distributions de probabilité au fur et à mesure que les élèves chargent les couleurs des billes.

Feuille de calcul pour représentation de l'évolution des distributions de probabilité

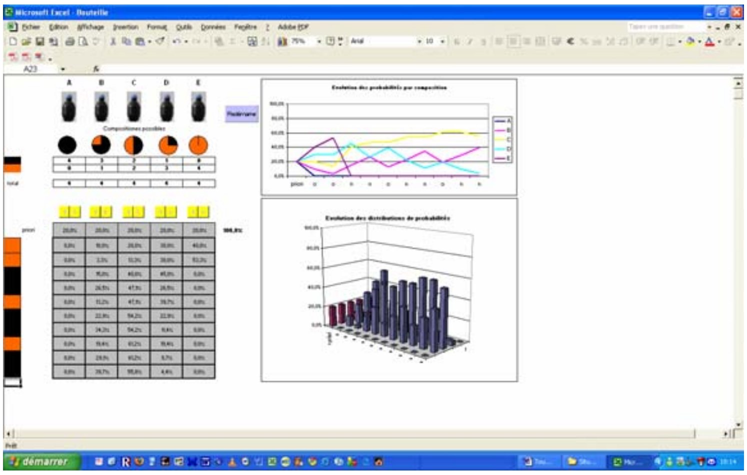


Figure 44

Ce passage à l'ordinateur n'implique pas l'introduction d'une simulation. L'objet produisant les événements aléatoires (les couleurs lors des retournements) reste toujours la bouteille, la seule fonction de l'ordinateur est de dispenser les élèves des lourds calculs des dix itérations.

Nous introduisons en plus un questionnaire par binôme où les élèves doivent interpréter les données de la feuille pour répondre aux questions. Dans ce questionnaire de vingt questions nous nous intéressons à l'utilisation du critère décisionnel de probabilité maximale et à celle de l'abduction comme raisonnement permettant de réassigner les probabilités. Dans cette étape de travail sur l'ordinateur, chaque binôme remplit un seul formulaire.

Cinquième étape : Institutionnalisation

Lorsque le questionnaire a été complété, l'enseignant réalise un tour de table visant, à la fin, une mise en commun des notions relatives à la probabilité bayésienne, au critère de probabilité maximale, au principe de raison insuffisante et aux principes du raisonnement par abduction. Par la suite nous allons effectuer l'analyse *a posteriori* de l'expérimentation de ce problème.

Analyse *a posteriori*

La salle et son équipement

Nous avons expérimenté cette troisième situation-problème le 5 Mars 2007. A cette occasion, dix sept des vingt deux élèves étaient présents. La est une représentation de la distribution des élèves et du matériel utilisé pour la séance. La salle disposait d'un vidéo projecteur (1) et de vingt et un ordinateurs (2). Nous avons placé une camera (3) au fond de la salle et quatre dictaphones (4) sur les tables des élèves dont un n'a pas fonctionné.

Pour cette séance nous avions prévu, en principe, un travail en binôme, mais la configuration de la salle a fait que les élèves discutent les uns par binôme, les autres par trinôme, les élèves travaillent donc en groupes de deux et trois.

D'ailleurs, les feuilles recueillies ne correspondent à une par groupe, ils nous est arrivé de trouver deux feuilles pour un même groupe d'élèves ou des groupes qui ne nous ont pas fait parvenir leurs feuilles de recherche. En tout, les élèves nous ont rendu quatorze feuilles. En retenant celles qui nous semblent appartenir à des groupes différents, le nombre se réduit à neuf. Nous effectuerons nos analyses à partir de trois types de sources: les neuf feuilles, les décryptages des enregistrements des trois dictaphones ayant fonctionné et l'enregistrement vidéo.

Distribution d'élèves et matériel. Séance 5 Mars

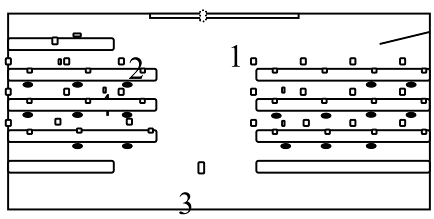


Figure 45

Le **Graphique 53** est une représentation de la distribution du temps de parole pour cette séance. Les groupes E, C et P correspondent à ceux des séances précédentes :

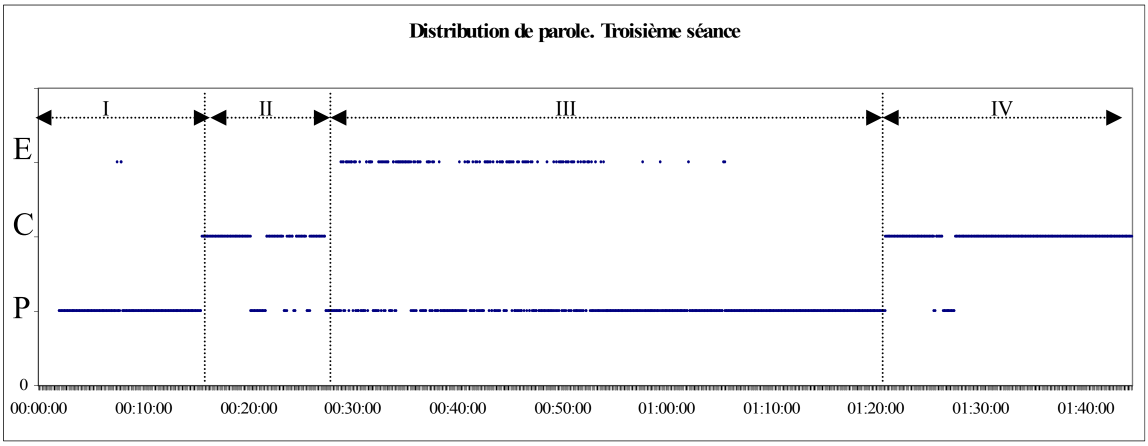
- Groupe P : L'émetteur du message est l'enseignant et il s'adresse à toute la classe.

- Groupe E : L'émetteur du message est un élève et il s'adresse à toute la classe.
- Groupe C : L'émetteur (enseignant ou élève) s'adresse à un membre de la classe ou à un groupe.

Sur ce graphique nous avons repéré les quatre premières étapes de notre analyse *a priori*. La dernière, celle de l'institutionnalisation, n'a pas eu lieu, au moins de manière formelle. Nous commençons l'analyse par l'introduction du problème.

Présentation du problème

Pendant les quinze premières minutes, l'enseignant explique le problème. Après avoir rempli l'urne de billes noires et orange, il en extrait quatre et les introduit dans la bouteille. Ce choix, nous l'avons ainsi convenu, se réalise au hasard et à la vue des élèves. De cette manière, nous évitons que les élèves perçoivent l'extraction des billes comme étant arbitraire. Si cela se produisait, les élèves chercheraient, à juste titre, de modéliser leurs croyances sur le choix de l'enseignant.



Graphique 53

Les trois distributions des jetons

Apres avoir expliqué les consignes et aidé les élèves à identifier les cinq hypothèses possibles, l'enseignant demande d'effectuer la première distribution de jetons (II, **Graphique 53**) :

1 *Voilà les règles dans la première colonne ma situation de départ c'est celle-là je vous demande*
5 *alors j'ai pas encore touché la bouteille je ne sais pas ce qu'il y a dedans vous non plus je vous demande d'attribuer dans la première ligne (…) le nombre de pièces que vous pensez pour chacune de compositions possibles (…)*

Première distribution de jetons

Pour leurs premières distributions, les élèves se servent de trois modèles. La **Figure 46** représente les types de distributions utilisées. Un seul groupe choisit la distribution à gauche (M), six groupes la distribution pyramidal (P) et deux autres l'equirépartie (E).

Types de distributions *a priori* employées

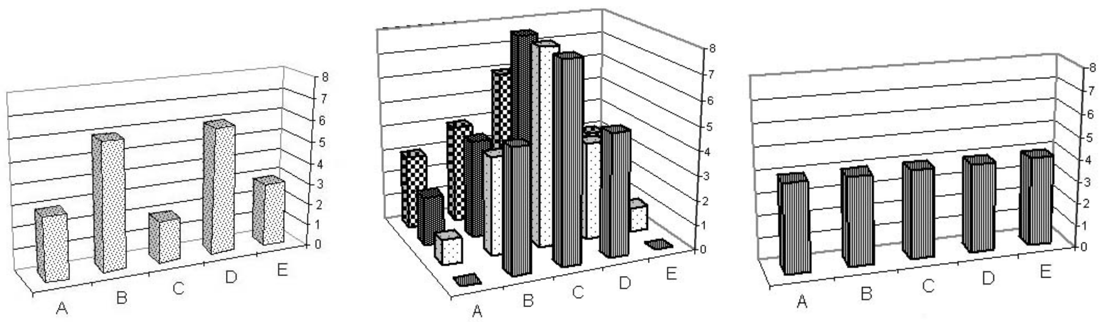


Figure 46

Les raisons exposées pour les élèves pour justifier leurs distributions sont variées et pas toujours faciles à déceler. Des trois groupes enregistrés par les dictaphones, un seul a choisi la distribution equirépartie (E), les deux autres la distribution pyramidale (P). Le décryptage correspondant au groupe de la distribution E nous renseigne sur un choix très rapide :

17 **Ne sachant pas combien il y a de boules noires ni de boules oranges le truc est équiprobable**
(…)
Quoi
Toutes les compositions sont équiprobables
On voit pas la différence entre entre toutes les possibilités
Ouais je suis d'accord

L'autre groupe ayant choisi aussi la distribution E donne des arguments proches, cette fois-ci nous les trouvons sur leur feuille : « **Les proportions sont équiprobables puisque nous n'avons aucune idée du nombre de boules noire et orange** ». Ces deux groupes utilisent de manière assez directe le principe de raison insuffisante pour leurs premières distributions *a priori*.

De leur part, les deux groupes enregistrés ayant choisi la distribution P, ne parviennent pas à un accord aussi immédiat. Par exemple, au sein d'un de ces groupes nous constatons deux positions différentes :

- D'une part un élève, pour qui les jetons représentent une somme à miser. En prenant les jetons comme un pari, cet élève reconnaît le risque associé. Il sait que les compositions monochromatiques sont possibles, mais comme pour lui elles sont peu probables, il prend le risque de ne rien leur assigner (« **On met dix au milieu et cinq aux autres** »). Les jetons ne représentent pas, pour lui, une mesure de certitude, mais un pari suivi d'un risque. En effet, pour lui, les hypothèses A (NNNN) et E (OOOO) sont très peu probables, l'argument pour leur attribuer une très faible probabilité étant la supposition d'une urne composée de cinquante pour cent de billes de chaque couleur, il décide de courir le risque de ne rien miser sur elles.

- D'autre part, un autre élève, pour qui les jetons représentent plutôt la notion de probabilité (« **Mais je mettrais un quand même çà peut sortir quand même (…) là tu les écartes** »). Cet élève insiste sur la possibilité que la bouteille soit du type A ou E et qu'il faut pour autant les attribuer des jetons, même s'il est d'accord sur leur faible probabilité.

Le sens des jetons est donc différent pour ces élèves, pour le premier il s'agit d'un pari où les croyances sont accompagnées d'un acte décisionnel. Pour le deuxième élève, les jetons représentent ses degrés de certitude sur les compositions possibles.

Les deux s'accordent sur le fait que la composition la plus probable est la C (NNOO), l'argument étant l'équirépartition supposée de billes noire et orange dans l'urne. La principale différence entre ces deux élèves réside dans le sens attribué aux jetons : une somme pour parier pour l'un (« *Çà c'est comme le loto tu prends des risques* »), une mesure de la possibilité d'une hypothèse pour l'autre. Finalement cette dernière approche prévaudra sur le pari dans ce groupe, et la distribution retenue est finalement 3-4-6-4-3.

Un autre groupe aussi, cette fois-ci non enregistré sur les dictaphones, interprète les jetons comme un pari, sa distribution n'attribue aucun jeton aux hypothèses monochromatiques (0-5-10-5-0), ils basent leur choix encore une fois sur la rareté de ces hypothèses.

Le troisième groupe enregistré choisit aussi la distribution P, néanmoins l'hypothèse d'une urne composée de cinquante pour cent de billes noires et cinquante pour cent d'orange n'est pas partagée pour tous les élèves de ce groupe. Pour un en particulier cette hypothèse n'est pas suffisamment solide. Toutefois il arrive à la même distribution par un autre argument, l'entropie de la série : « *Il est plus probable d'obtenir de la diversité malgré le fait que nous ne connaissons pas les proportions de couleurs* »). De cette manière, pour cet élève, moins monotone est une série, plus probable elle est.

En résumé, les groupes ayant choisi la distribution pyramidale (P) justifient leur répartition par deux types d'arguments, l'un l'admission d'une distribution équirépartie dans l'urne, l'autre par l'entropie des compositions, il arrive aussi que quelques élèves interprètent les jetons en terme de pari.

Pour ce qui concerne le seul groupe ayant choisi la distribution à gauche (M) (**Figure 46**), leurs raisons ne sont pas évidentes à interpréter. Nous disposons de la réponse rédigée sur la feuille et des explications fournies lors de l'étape III où s'effectue le bilan de cette partie. Tant la réponse sur la feuille comme l'explication de l'élève dans l'étape III nous semblent contradictoires, probablement ce groupe a interprété le nombre de jetons comme un pari. Le **Tableau 62** résume les justifications par type de distribution.

Justifications des élèves pour les distributions *a priori* employées

	Type de distribution		
	M	P	E
Justification	*Pari*	*Urne equierpartie, entropie de la composition et pari*	*Raison insuffisante*

Tableau 62

Deuxième distribution de jetons

Ayant rempli les élèves la première ligne de leur tableau, l'enseignant retourne la bouteille, la couleur obtenue est orange (min. 20 de la séance). Ensuite il demande d'analyser la possibilité d'une révision des croyances à partir de cette couleur observée. Les élèves commencent donc à discuter sur les possibles modifications de leurs distributions.

Tout de suite les élèves arrivent à la conclusion déterministe : la bille orange fait que la composition A (NNNN) devient impossible. Néanmoins, la redistribution sur les autres compositions ne fait pas l'unanimité entre les groupes. Par exemple, les élèves ayant choisi la distribution P réagissent de manière différente en ce qui concerne les pièces à replacer :

- Un groupe n'enlève que les jetons de la composition devenue impossible (A) et les distribue de manière équitable sur les trois compositions les plus probables (C, D et E) (**Tableau 64**).

	A	B	C	D	E
Distribution *a priori*	3	4	6	4	3
Distribution *a posteriori*	0	4	7	5	4

Tableau 63

- Un autre groupe enlève en plus un jeton de la composition B (NNNO). De cette manière quatre jetons sont à replacer (**Tableau 64**).

	A	B	C	D	E
Distribution *a priori*	3	4	6	4	3
Distribution *a posteriori*	0	3	7	6	4

Tableau 64

Pour le premier groupe la réassignation fut équitable. Pour le deuxième, la mobilisation du raisonnement abductif semble plus nette, deux éléments le confirment : d'une part l'action d'enlever un jeton de B et d'autre part l'assignation de deux nouveaux à D. En effet, une bille orange affaiblit la possibilité de l'hypothèse B (NNNO) en même temps qu'elle renforce celle de l'hypothèse D (NOOO). L'élève aurait suivi une démarche abductive pour déterminer sa distribution *a posteriori.*

Néanmoins, cet élève semble s'être servi d'un autre argument qui aurait agi de manière complémentaire au celui de l'abduction. Cet autre argument, que nous avons dénommé comme entropique, consiste à croire qu'une hypothèse monochromatique est toujours peu probable. De cette manière, le raisonnement par abduction se combine avec cette conception menant l'élève a assigner un seul nouveau jeton à l'hypothèse E.

Un autre exemple de la conjonction de l'argument abductif et de l'entropique peut se trouver dans les distribution d'un autre groupe où l'abduction s'applique clairement sur l'hypothèse B et très faiblement sur l'hypothèses E (**Tableau 65**):

	A	B	C	D	E
Distribution a priori	0	5	10	5	5
Distribution a posteriori	0	1	10	8	1

Tableau 65

Néanmoins, d'autres arguments semblent se manifester en amont de l'abduction et de l'entropie de la composition. Un troisième groupe décide d'effectuer des déplacements de manière telle à que la distribution *a posteriori* reste symétrique. Cela pourrait donc constituer un autre argument qui s'intègre aux précédents. Non seulement les déplacements se réaliseraient par des arguments abductifs ou entropiques, mais aussi par une raison géométrique qui consisterait à chercher une distribution ayant la propriété de symétrie axiale.

Des distributions a *posteriori* symétriques, nous en avons trouvés chez trois groupes en tout, un exemple de ce genre de distributions est reproduit dans **Tableau 66**).

Distributions symétriques a priori et a posteriori

	A	B	C	D	E
Distribution a priori	2	4	8	4	2
Distribution a posteriori	0	4	6	6	4

Tableau 66

Pour ce qui concerne les deux équipes ayant choisi comme première distribution l'équirépartie (4-4-4-4), leurs conclusions tirées de la couleur de la bille restent exclusivement sur le plan déductif. En effet, les deux groupes ayant basé leur première distribution sur le principe de raison insuffisante retiennent leur distribution équirépartie (0-5-5-5-5) (« **A n'est plus possible car on sait qu'il y a au moins une boule orange** »). Ces élèves restent attachés à l'argument de ce principe et continuent à affirmer ne rien connaître sur la composition de la bouteille, sauf qu'elle n'est pas du type A.

Finalement, le groupe ayant choisi la distribution M (3-6-2-6-3) réalise une inversion, il passe d'une distribution où la composition C (NNOO) est la moins probable à une autre où elle devient celle de plus haute probabilité (0-6-8-6-0). Le sens attribué aux jetons semble pour ce groupe se confirmer comme étant d'un pari, la composition E étant possible, les élèves décident de n'en rien lui assigner (« *Il y a une très faible chance pour n'avoir que de billes oranges ainsi que noires la composition C a beaucoup plus de chances d'être tirée par proportionnalité* »).

Troisièmement distribution de jetons

A la minute vingt cinq, l'enseignant retourne pour la deuxième fois la bouteille, il obtient une bille de couleur noire. Les élèves écartent facilement la composition E (OOOO).

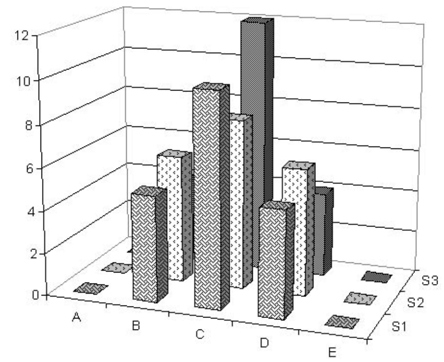
Les distributions sont maintenant très proches et répondent toutes au type pyramidal. Toutes se caractérisent par représenter la composition C (NNOO) comme étant l'hypothèse la plus probable, même pour les deux groupes qui avaient représenté une distribution équirépartie (E) lors des deux premières distributions.

Ces derniers élèves, rappelons-le, ont choisi comme première distribution *a posteriori*, une équi-répartie. De cette manière, ils n'ont abordé que les conclusions déterministes du premier retournement. Après le deuxième retournement, ils choisissent une distribution de forme pyramidale, ce qui pourrait nous mener à penser que ces élèves auraient quitté le stade déterministe pour rentrer dans l'indéterministe, néanmoins dans une des feuilles de ces deux groupes nous trouvons des traces de la cohabitation des arguments tant abductifs que déductifs :

> *« E n'est plus possible et nous mettons des valeurs équiprobables au 3 restant mais il faut mettre les 20 pièces alors les 2 derniers jetons on été sur la C car les tirages ont donné autant de noir que d'orange comme en C »*

Le **Graphique 54** montre les différentes distributions *a posteriori* utilisées par les élèves après le deuxième retournement de la bouteille. Les répartitions sont toutes symétriques et concentrées sur la composition C (NNOO). Pour les justifier, sauf le cas ci-dessus commenté, les élèves se limitent à expliquer l'impossibilité de la composition E (OOOO) due à l'apparition d'une bille noire dans le dernier retournement.

Dernières distributions *a posteriori*



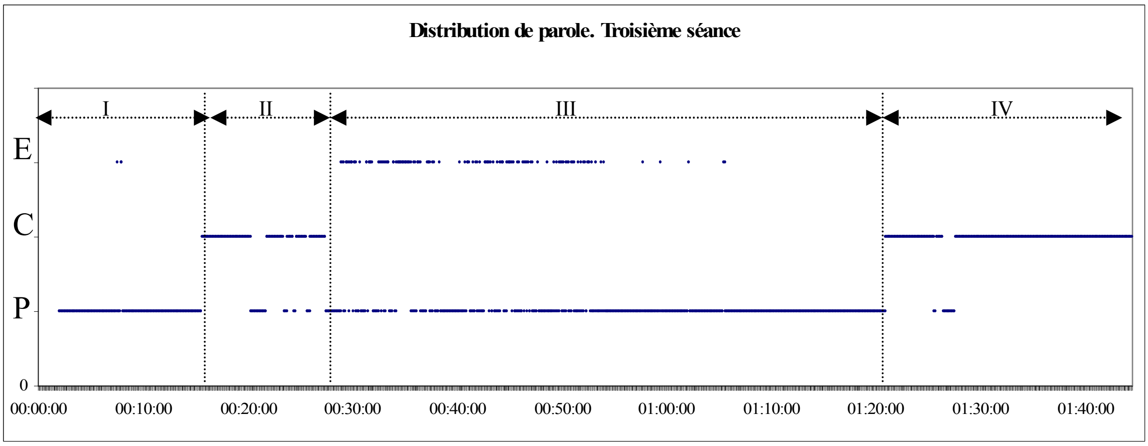
Graphique 54

En fait, les explications des élèves sont plus précises et riches lorsqu'elles sont guidées par l'enseignant. Par la suite, nous analyserons l'étape de mise en commun où la classe discute les distributions choisies et les justifications.

Premier bilan et révision de distributions

Cette étape (III, **Graphique 55**) commence à la minute vingt sept de la séance et se termine à la minute quatre vingt, le graphique est une représentation par la répartition du temps de parole où deux grands moments peuvent être identifiés. Le premier, caractérisée par une forte participation des élèves (E), le deuxième par une monopolisation progressive de la parole de la part de l'enseignant.

Dans la première partie de cette troisième étape se réalise le bilan du travail avec les jetons, les élèves y expliquent leurs distributions et développent leurs justifications. Dans cette première partie aussi, l'enseignant effectue le respectif bilan. La deuxième partie est consacrée à l'explication de la formule de Bayes. Nous commençons par analyser la première partie.



Graphique 55

Bilan des distributions des jetons

A la minute vingt sept l'enseignant interroge la classe sur la première distribution de jetons, celle qui précède le premier retournement de la bouteille. Un premier groupe d'élèves répond avoir choisi la suite (2, 4, 8, 4, 2). Sa justification pourrait se résumer comme étant une adaptation du principe de raison insuffisante, appliqué non pas sur les compositions possibles mais sur les couleurs possibles. En d'autres termes, ces élèves admettent ne rien savoir sur la composition de l'urne, ce manque de connaissance conduit ce groupe à attribuer la même chance à chaque couleur (min 29) : « **On a autant de chances d'avoir une boule orange et une boule noire on peut pas le savoir donc on a une chance sur deux** ».

Cette hypothèse sur les couleurs permet à ce groupe de trouver un ensemble de référence pour évaluer les probabilités de chaque composition possible. Cette évaluation se réalise de manière intuitive, ils assignent plus de jetons aux compositions centrales et moins aux extrêmes.

Un groupe conteste ces arguments, ces questions s'adressent à l'hypothèse d'équi-répartition de l'urne : « **Je ne sais pas s'il y a cinquante pour cent de chances de tomber sur une boule noire et orange** ». Cette réflexion provient d'un des deux groupes ayant choisi la distribution (E). Ces deux groupes, rappelons-le, ont choisi une distribution de probabilités sur les compositions du type équi-répartie, leur justification se basait sur le principe de raison insuffisante.

Plusieurs présentations se succèdent, et progressivement deux approches se dessinent pour évaluer les probabilités :

- d'une part, ceux qui accordent à l'urne une distribution équi-répartie. Cette hypothèse permet à ces groupes de trouver un ensemble de référence fini. L'évaluation de la probabilité de chaque composition s'effectue de manière intuitive à partir de cet ensemble de référence. Leur distribution de probabilités est du type pyramidal (P).
- D'autre part, ceux qui attribuent la même probabilité à chaque composition avec l'argument de ne rien savoir sur la composition de l'urne. Pour ces élèves il n'y a aucun ensemble de référence.

La plupart des groupes ont choisi une distribution pyramidale. Deux groupes ont choisi la distribution uniforme. Ces derniers groupes contestent chaque présentation du type pyramidal, leurs arguments concernent l'impossibilité d'assurer l'équi-répartition des couleurs dans l'urne, voici quelques extraits de ces réfutations :

> « **En fait on ne connaît pas le contenu de l'urne** », « **théoriquement on a autant de chances au départ d'avoir les cinq possibilités** », « **il y a autant de chances de probabilité de tomber sur ces possibilités** », « **Mais on ne peut pas dire qu'il y a cinquante pour cent de noires et cinquante pour cent d'oranges** », « **C'est pareil on ne connaît pas les couleurs qu'on a dans la bouteille** », « **Comme on ne sait rien il peut il avoir dix pour cent de noires et quatre vingt dix pour cent d'orange pour moi c'est toutes équiprobables** ».

Le principe de raison insuffisante

Ces réflexions ne sont jamais contestées par la classe, aucun groupe ne fournit des arguments en faveur de l'hypothèse d'une équi-répartition de couleurs dans l'urne. Les présentations des groupes ayant choisi une distribution pyramidale répondent à un cycle, d'abord un groupe présente la distribution, puis la justification par l'urne équi-répartie et finalement la réfutation provenant des deux groupes de la distribution E.

L'enseignant intervient dans un de ces cycles où les groupes ne parviennent pas à résoudre leurs différences. Dans son intervention, l'enseignant valide l'argument du principe de raison insuffisante (min. 40) :

4 *je ne sais rien sur cette urne donc (…) je ne peux pas prendre position (…) je ne peux pas*
0 *privilégier a priori comme je ne sais rien a priori plus une composition qu'une autre (…) et puis comme je ne sais strictement rien et bien qu'est-ce qu'on a comme réponse*
Quatre quatre quatre quatre et quatre

Ensuite la classe continue avec la première distribution *a posteriori*, celle qui suit à la bille de couleur orange. La conclusion déductive tirée du premier retournement est rapidement donnée : « **On peut barrer NNNN** ». Néanmoins, les conclusions abductives n'émergent pas facilement, elles devront attendre l'intervention de l'enseignant pour être discutées par la classe.

Les deux groupes ayant proposé la principe de raison insuffisante parviennent facilement à imposer leur distribution *a posteriori* (0-5-5-5-5). Les élèves expliquent que la situation est similaire à celle qui précédait le retournement : on ne connaît rien sur la composition de la bouteille sauf que cette fois-ci à la place de cinq, ils considèrent quatre hypothèses possibles (B, C, D et E).

L'approche soutenue par ces deux groupes pourrait se caractériser comme étant déterministe. Ils ne tirent que des conclusions déductives : une bille orange implique que l'hypothèse A (NNNN) est impossible. Ces élèves n'effectuent pas des inférences abductives à partir de la couleur de la bille observée. Par exemple, une approche abductive permettrait d'inférer qu'une bille de couleur orange renforce les hypothèses où la proportion d'orange est plus importante.

En particulier, une première bille orange ferait que la composition E (OOOO) devienne la plus probable. En effet, le phénomène observé (O) est plus attendu sous cette hypothèse que sous les autres. De cette manière, l'hypothèse E explique mieux que les autres la bille observée, il est pour autant plus probable que la bouteille soit du type E. Ce genre d'inférences n'est pas pris en compte par ce groupe qui déclare ne rien savoir sur la composition sauf qu'elle n'est du type A (NNNN).

Le raisonnement par abduction

L'enseignant, pou sa part, en ne trouvant pas des réactions négatives à la distribution donnée (0-5-5-5-5) continue avec le deuxième retournement : une bille de couleur noire. L'approche déterministe continue à s'imposer et la distribution proposée est toujours équirépartie (0-6,6-6,6-6,6-0). L'argument reste toujours du type déductif, il s'agit d'écarter la composition E (OOOO) comme conséquence de la bille noire obtenue.

L'enseignant se voit obligé d'intervenir pour faciliter l'émergence du raisonnement par abduction (min. 42) :

4 (…) *bon alors réfléchissons un petit peu* (…) *vous avez tiré une fois orange et vous avez tiré une*
2 *fois noire et vous me dites que ayant tiré une fois orange et une fois noire vous avez la même degré de certitude d'obtenir celle-là que celle-là que celle-là est-ce que vous êtes d'accord avec çà*

Cette réflexion de l'enseignant semble fonctionner comme une approbation donnée à ceux qui, lors de l'étape de travail en groupe, avaient réfléchi à partir d'une approche abductive. En effet, cette réflexion de l'enseignant suffit pour que la classe commence à proposer des conclusions abductives :

En fait comme on a tiré une fois orange et une fois noire on pourrait dire que çà peut être cinquante pour cent qu'il y a des boules oranges et de boules noires dans ce cas on favorise la composition C alors

L'enseignant accompagne cette première réflexion et l'élève termine par cerner son approche en proposant une distribution pyramidale (0-6-8-6-0). Ce type de distribution est de plus en plus partagée et acceptée par toute la classe.

Il semble que l'approche adductive se trouvait latente chez ces élèves et qu'il aura suffi d'un signe de la part de l'enseignant pour que la classe déploie des arguments probabilistes. En quelques minutes, la classe parvient à se mettre d'accord sur la composition B (NNOO) comme plus probable que les autres. En effet, progressivement, et en se servant de comparateurs du genre « plus que », « moins que » et « autant que », les élèves précisent leurs distributions, toutes de type pyramidal.

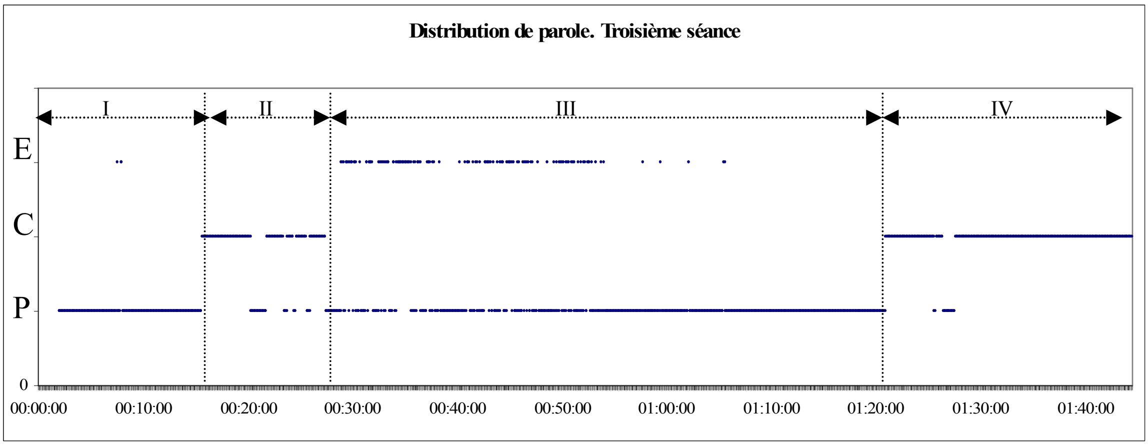
Ensuite, l'enseignant propose d'autres issues possibles pour les deux retournements de la bouteille. Il demande aux élèves d'indiquer leurs distributions de jetons s'ils avaient obtenu deux billes orange au lieu d'une orange et une noire. A cette nouvelle question, les élèves répondent à partir de l'approche abductive :

De favoriser D et E de baisser B et C…
Non de baisser B et d'augmenter C bah C D E quoi
4 Parce que s'il y a deux boules oranges différentes qu'on a eu B il vaut normalement zéro non
7 mais on n'est pas sûr que ces soient différentes parce que peut être la même donc B est çà peu arriver quand même on laisse un peu à B

Dans tout ce temps consacré aux principes du raisonnement par abduction (min. 43 à 53), l'enseignant insiste sur les deux étapes caractéristiques à cette démarche, l'une qui établit la mesure de l'attente d'un phénomène sous une hypothèse donnée ($P_{Hj}(O)$), l'autre concernant le procédé abductif proprement dit, où il s'agit d'estimer la possibilité d'une hypothèse lorsque le phénomène est constaté ($P_O(Hj)$). Par exemple (min. 49), pour se référer à ($P_O(E)$) l'enseignant s'exprime en termes de : « **le fait d'avoir orange çà a plutôt tendance à me conforter dans l'idée que [la bouteille est de type E(OOOO)]** ». Et lorsqu'il souhaite se référer à ($P_B(O)$) , il parle de combien on attendrait d'avoir orange si la bouteille était du type B. Pour se référer à la modification d'une probabilité, enseignant et élèves parlent de « faire monter » ou « faire baisser ». Ces idées semblent comprises pour les élèves qui s'expriment par exemple en termes de « favoriser » ou « augmenter » pour signifier qu'une composition devient plus probable en fonction de la bille observée (min 45. «**Je favoriserais aussi la C parce que vue que les tirages** »).

A la minute cinquante cinq, l'enseignant introduit les représentations symboliques et numériques de la probabilité. Sur le registre langagier, le terme probabilité est apparu pour la première fois à la minute dix huit, et il a été couramment utilisé pour se référer aux jetons. Néanmoins, c'est la première fois que l'enseignant le représente sur les registres numérique et symbolique. Pour introduire la représentation numérique, il demande aux élèves d'indiquer en probabilités la première distribution de jetons. Les élèves répondent en termes de chances et sous la forme de fractions (min. 54 « **une chance sur cinq** »), c'est l'enseignant qui ensuite introduit l'expression « un cinquième ».

Le passage aux registres symboliques et numériques marque le début de la deuxième partie de la troisième étape (III, **Graphique 56**). Cette partie se caractérise par une forte participation de l'enseignant. Dans ces vingt minutes qui suivent, l'enseignant essaie de transposer au registre symbolique la démarche abductive discutée précédemment. L'expression sur le registre symbolique (formule de Bayes) devrait, en théorie, représenter un moyen permettant aux élèves de déterminer avec précision ce qu'ils avaient fait de manière intuitive.



Graphique 56

Pour évaluer la probabilité *a posteriori* pour une hypothèse donnée (Hj), la formule de Bayes considère le quotient entre l'attente du phénomène observé sous une hypothèse donnée (P_{Hj}(O)) et son attente moyenne (P(O)). Le premier étant fixe, le deuxième est en fonction des probabilités des hypothèses. Ce quotient devient taux de changement pondérant la probabilité *a priori* (P(Hj)).

$$P_O(H_j) = \frac{P_{Hj}(O)}{P(O)} \times P(H_j)$$

Dans cette séance (P(Hj)) semble bien comprise, son signifié a été présent dès le début de la séance, de même que (P_{Hj}(O)), même si les élèves ne sont pas habituellement confrontés à ce genre d'écriture. Néanmoins, P(O) en tant qu'attente moyenne (probabilité totale) peut être considérée comme un concept nouveau pour ces élèves. En effet, cet objet symbolique est généralement présenté dans les exercices comme étant un élément de la formule de Bayes, en d'autres termes, dépourvu de sens, rares seraient les occasions où les élèves auraient été confrontés à des exercices où la probabilité totale est porteuse de sens.

La formule de Bayes

Pour commencer à introduire la formule de Bayes dans cette séance, l'enseignant choisit d'interroger les élèves sur le dénominateur de la formule, P(O). Ce calcul, n'étant pas directement nécessaire à la résolution du problème, c'est l'enseignant qui l'introduit.

Il y aurait en principe deux grandes voies pour s'intéresser à ce calcul, l'une qui interroge, tel que l'a fait l'enseignant, à partir d'un de ses signifiants (min. 56 « **est-ce qu'on peut trouver la probabilité d'O** »), l'autre qui le fait à partir de son signifié (par exemple :

pourriez vous déterminer en moyenne, combien on attend une bille orange ? »). Aborder la question à partir de son signifié ou de son signifiant pourrait ne pas être équivalent pour les élèves. La première option pourrait être plus coûteuse pour les élèves. En effet, cette question posée dans une situation-problème contextualisée requerrait un travail additionnel de construction de son signifié, pour ensuite repérer les éléments du problème nécessaires à sa réponse.

La construction du sens de la question pourrait ne pas être la seule difficulté. En général, les élèves sont confrontés à des situations de calcul de probabilités bien différentes de celle proposée ici par l'enseignant. Pour eux, les calculs de probabilités d'événements s'effectueraient lorsque les hypothèses sont connues, le schéma pourrait se résumer comme étant un chemin allant des hypothèses vers les données, les premières sont connues, et fournies par les énoncés des exercices, les deuxième sont à la charge de l'élève. Selon ce schéma, sans l'hypothèse, on ne pourrait pas déterminer la probabilité d'un événement. Il semble que les élèves se sont basés sur ce genre d'exercices pour répondre qu'il est impossible de déterminer la valeur de P(O) (min. 57).

Devant cette réponse, l'enseignant décide d'aborder la question autrement, il apporte des éléments nécessaires à son calcul : ($P_A(O)$, $P_B(O)$, …, $P_E(O)$) ; et même, en faisant recours à la mémoire didactique de la classe, il dessine un arbre de probabilités sur le tableau où les élèves doivent remplir les valeurs des probabilités de chaque branche. Néanmoins, ces éléments ne suffisent pas pour faire émerger la formule de la probabilité totale travaillée lors de séances précédentes.

En fait, il devient difficile aux élèves d'associer l'arbre de probabilités au problème de la bouteille, ils ne parviennent pas à déterminer l'information correspondant à chaque branche de l'arbre. De plus, de "nouvelles" expressions sont utilisées par l'enseignant dans le registre langagier dont ils ont du mal à trouver la représentation symbolique (« P de O sachant A », etc.). Enfin, la participation des élèves décline fortement et l'enseignant prend à sa charge la réponse à la question de la valeur de la probabilité d'obtenir une bille orange.

Dans les minutes que suivent, les interventions des élèves se restreignent à résoudre des tâches calculatoires demandées par l'enseignant. Celui-ci développe complètement la démonstration de la formule de Bayes. Cette tâche, est suivie de la détermination de la première distribution *a posteriori* et de son interprétation. En résumé, la démonstration de la formule, le calcul des probabilités a posteriori et leurs interprétations l'occupent de la minute soixante dix à la quatre vingt.. L'exposé finalement est suivi d'une pause de cinq minutes.

Le développement de la formule de Bayes semble répondre à une habitude de l'enseignant ou du métier d'enseignant de mathématiques. En effet, dans un entretien non enregistré à la fin de la séance, ils nous confiera que l'introduction de la formule sans sa démonstration, lui semblait une présentation incomplète et que, bien que difficile à suivre cette démonstration, en la développant, les élèves sauraient d'où provient l'expression de Bayes. Quelques minutes plus tard, la séance finie, il nous confiera aussi que probablement la démonstration effectuée de la formule pouvait avoir répondu plutôt à une habitude professionnelle qu'à une nécessité des élèves. Pour lui, en tant qu'enseignant de mathématiques, une présentation serait plus propre si elle est accompagnée de sa démonstration.

Travail sur l'ordinateur

A la minute quatre vingt et une, et après cinq minutes de pause, l'enseignant présente la nouvelle étape où les élèves travaillent sur l'ordinateur (IV, **Graphique 56**). Après avoir distribué un questionnaire par binôme, l'enseignant explique le fonctionnement de la feuille de calcul. La **Figure 47** est une représentation de l'interface tableur sur laquelle travaillent les élèves dans cette dernière étape de la séance.

La feuille de calcul comporte trois sections :

- Une première où sont visualisées les différentes compositions possibles et les proportions de billes noire et orange.
- Une deuxième où les élèves interagissent avec l'interface. Cinq boutons jaunes permettent de déterminer les premières probabilités *a priori*. En cliquant sur ces boutons, la distribution apparaît sur la première ligne grise du tableau. Lorsque la somme de toutes les probabilités est égale à l'unité, le rectangle rouge disparaît et l'algorithme s'active, en cas contraire il reste bloqué. La colonne vert clair à gauche permet d'introduire la couleur de la bille observée (« n » ou « N » pour une bille noire, « o » ou « O » pour une orange. Une fois la couleur confirmée (en tapant sur la touche « enter »), la cellule prend la couleur de la bille indiquée. La distribution *a posteriori* est automatiquement calculée par l'algorithme de Bayes.
- Une troisième section où les distributions sont représentées graphiquement. Le graphique supérieur permet une visualisation en deux dimensions, l'inférieur en trois. Toutes les distributions de probabilités se représentent de manière automatique. La feuille étant protégée, la seule section interactive correspond à la première distribution

a priori (boutons jaunes) et aux couleurs des billes (colonne vert clair). Les autres sections de la feuille restent bloquées aux modifications. Finalement, un bouton « redémarrer » permet de remettre la feuille à zéro.

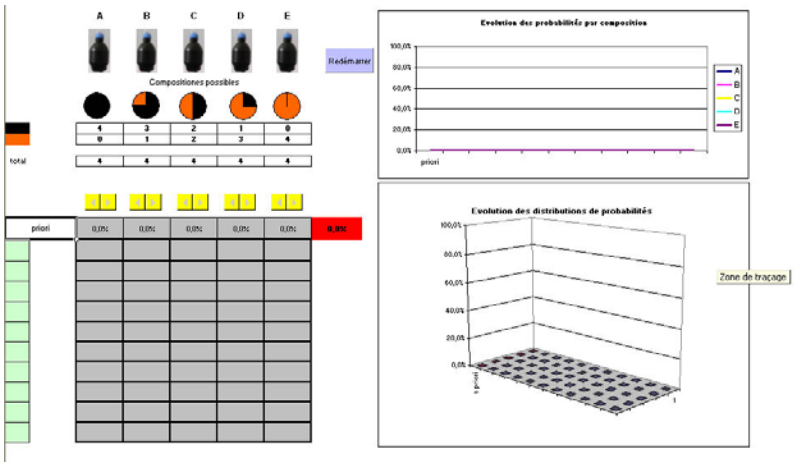


Figure 47

Le questionnaire accompagnant le travail sur l'ordinateur comporte vingt questions. Dans ce questionnaire nous nous intéressons à l'utilisation du critère décisionnel de probabilité maximale et à l'abduction comme raisonnement permettant de réassigner les probabilités. Toutes les questions sont posées dans le registre langagier, elles se caractérisent pour être les unes à choix multiple, les autres ouvertes. Les premières, pour choisir une réponse parmi un ensemble de possibles, par exemple une composition (NNNN, NNNO, NNOO, OOON, OOOO). Les ouvertes cherchent des justifications.

La transition

Dans l'étape précédente, et lorsque l'enseignant expliquait la formule de Bayes, les élèves étaient intervenu très rarement (III, **Graphique 56**). Ces interventions pourraient être caractérisées comme étant des tâches calculatoires (produits). L'enseignant guidait les réflexions, les élèves réalisaient les calculs sollicités. La dernière de ces tâches calculatoires avait concerné le calcul de probabilités du type $P(H_i \cap O)$. C'est cette notion de probabilité que quelques élèves cherchent à trouver en premier lieu sur la feuille de calcul. De cette manière, ces élèves cherchent une continuité entre les étapes.

Pour sa part, l'enseignant semble avoir choisi une voie personnalisée pour expliquer en détail cette nouvelle étape de la séance. Au lieu d'une présentation générale pour toute la classe, il visite les groupes et répond aux questions des élèves. Puis il revient sur la bouteille et effectue les huit retournements manquants. La suite complète des dix couleurs est O,N,N,O,N,O,O,N,O,O. La **Figure 48** représente la feuille de calcul pour cette suite de retournements.

Feuille de calcul pour série de retournements

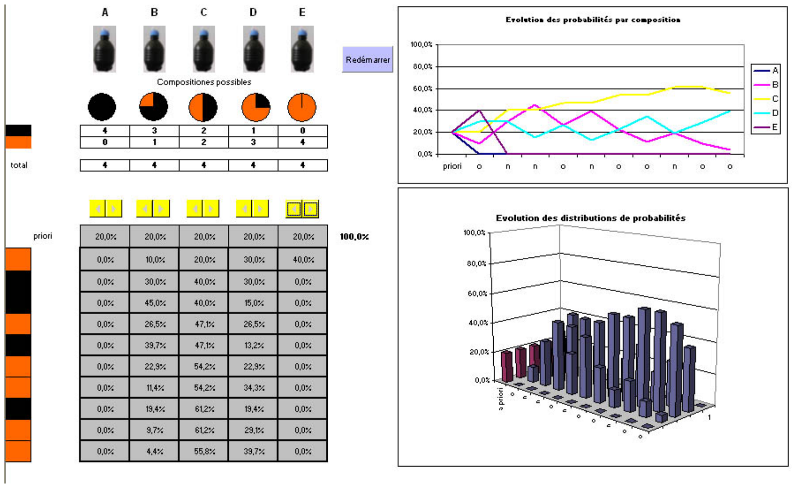


Figure 48

Cette suite donne une proportion d'oranges de 60 %. Quelques élèves appliquent intuitivement les principes de la méthode de maximum de vraisemblance et indiquent la composition C (NNOO) comme étant la plus probable. En effet, l'hypothèse qui maximise la probabilité de l'échantillon est effectivement la composition C. Ces élèves ne réfléchissent pas dans ces termes, pour eux, la composition C est la plus probable parce qu'elle est la plus proche de l'échantillon obtenu.

L'argument de la proximité entre échantillon et composition semble suffisant aux élèves et ils n'éprouvent pas la nécessité d'évaluer numériquement la situation en termes de probabilité. Ceci constitue, à nos yeux, une des grandes difficultés didactiques de la formule de Bayes. En d'autres termes, l'argument de la distance entre la proportion de l'échantillon est plus économique que la formule de Bayes. Il est toujours plus simple de justifier un choix par la proximité des deux proportions que par la probabilité maximale donnée pour la formule de Bayes.

D'ailleurs, la classe intègre progressivement la feuille de calcul au questionnaire et à la minute quatre vingt dix, les groupes travaillent sur leurs questionnaires. Des dix sept élèves présents dans de cette séance, nous disposons de quatorze questionnaires. Par la suite, nous analyserons les réponses des élèves à ces questionnaires.

Analyse de réponses des élèves au questionnaire

Parmi les quatorze questionnaires rendus, nous en avons identifié deux appartenant à un même groupe d'élèves. Nous avons décidé de les retenir dans ces analyses car leurs justifications sont parfois différentes. De cette manière, même si les réponses aux questions à choix multiples coïncidaient dans ces questionnaires, les réponses ouvertes nous fournissaient plus d'information qu'un seul questionnaire.

La première question, à choix multiple, concernait l'estimation de la composition de la bouteille après les dix retournements, ensuite nous avions demandé de justifier la composition choisie. Pour les quatorze questionnaires, la composition estimée est la C (NNOO). Néanmoins, à la question de savoir s'ils étaient persuadés de la composition, la moitié (sept réponses) répond oui, un élève ne répond pas et les six autres répondent ne pas être persuadés de cette composition (**Tableau 67**).

	Est-vous persuadé de la composition ?
Oui	7
Non	6
Sans réponse	1

Tableau 67

Pour ce qui concerne les raisons, elles sont variées. Parmi les élèves ayant répondu être persuadés de la composition, seuls cinq élèves donnent des justifications. Nous retiendrons trois réponses :

- l'urne serait composée de cinquante pour cent de billes orange et cinquante pour cent de billes noire. Avec cette composition, il est plus probable d'obtenir la composition C (NNOO).
- l'hypothèse C, après dix tirages, a 55,8 % de probabilité d'être la composition de la bouteille.
- La proportion de la composition C est la plus proche de celle de l'échantillon.

Les réponses nous suggèrent que ces élèves étaient persuadés du choix plutôt que de la composition. Pour ce qui concerne les élèves ayant répondu ne pas être persuadés de la composition (six sur quatorze). Nous retiendrons deux types de réponses :

- La proportion obtenue ne peut pas assurer la composition car les billes observées pourraient être les mêmes (trois réponses)
- L'hypothèse C a 55,8% de probabilité d'être la composition de la bouteille.

En analysant toutes les réponses à ces trois premières questions, nous observons que l'information fournie par la feuille a été citée à trois reprises. Ce faible taux pourrait s'expliquer par la non assimilation de l'information de la feuille. En effet, ces réponses correspondent aux premières questions, les élèves pourraient ne pas avoir déchiffrer encore l'information disponible sur l'ordinateur.

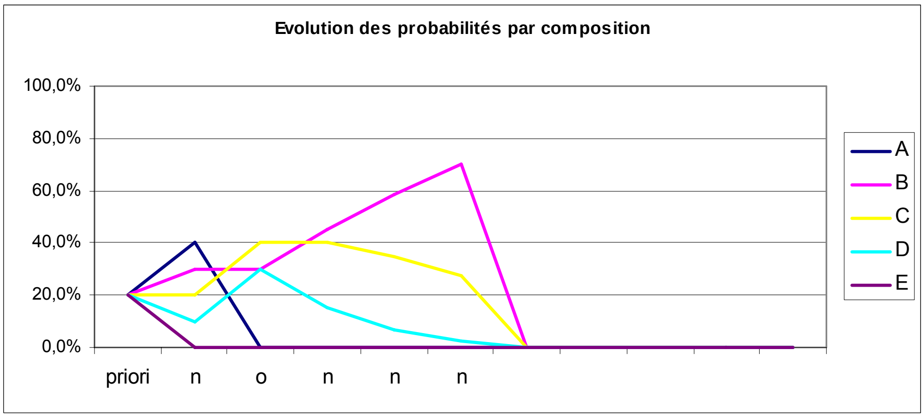
D'ailleurs, trois élèves ont répondu ne pas être sûr de la composition à cause de la répétition possible de la même bille lors des retournements. Ce type de réponse nous renseigne sur les difficultés des élèves à raisonner en termes probabilistes. Ce même argument à été mentionné dans cette séance lors des étapes précédentes. Nous continuons l'analyse du questionnaire.

La question suivante demande si le choix de composition aurait été le même si, à la place de dix, on n'avait fait que cinq retournements, les cinq premiers. Nous avions inclus cette question avec l'intention de motiver le débat sur le choix d'une composition. La probabilité d'une composition est fonction de deux paramètres, l'un la première distribution *a priori*, l'autre la proportion de couleurs des billes. Le premier était fixe pour cette séance, nous avions joué sur le deuxième. En principe, il est possible, qu'au bout de cinq retournements, la proportion de couleurs soit significativement différente de celle de dix. De cette manière, le choix lors de cinq retournements serait différent au celui de dix. Par exemple, si lors des cinq premiers retournements les couleurs sont N, O, N, N, N, la composition la plus probable sera la B (NNNO) (**Tableau 68** et **Graphique 57**)

Distribution de probabilités pour retournements NONNN

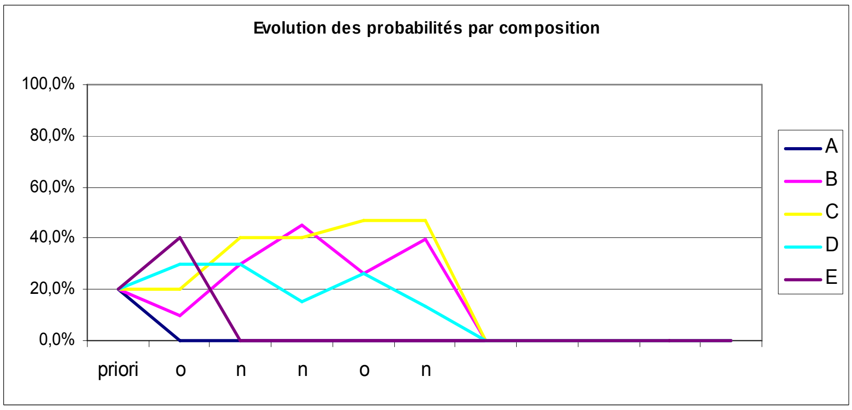
A (NNNN)	B (NNNO)	C (NNOO)	D (NOOO)	E (OOOO)
0	69,8%	27,6%	2,6%	0

Tableau 68



Graphique 57

Pour cette séance, les cinq premiers retournements (O, N, N, O, N) donnent une distribution de probabilité où la composition C (NNOO) est toujours la plus probable (**Graphique 58**).



Graphique 58

Dans les quatorze questionnaires, douze élèves ont choisi la garder la même composition pour cinq retournements (**Tableau 69**). Les justifications (neuf) nous suggèrent que la feuille est progressivement intégrée aux analyses des élèves, sept élèves se sont servis des distributions de probabilité pour choisir la composition C (« **Car lors du tirage N° 5 NNOO a 47,1 % c'est la plus important** »), pour ces élèves donc, le critère de choix est la probabilité maximale.

	Auriez-vous fait le même choix si on avait fait juste cinq retournements ?
Oui	12
Non	1
Sans réponse	1

Tableau 69

D'ailleurs, une caractéristique particulière de la notion bayésienne consiste en ce que la valeur de la probabilité n'est pas fixe. En principe cette valeur, en tant qu'appréciation personnelle, peut différer d'un observateur à l'autre et même elle peut évoluer par rapport à un même observateur. Le problème de la bouteille correspond à ce dernier cas, où les probabilités pour les compositions varient en fonction de l'information tirée des retournements successifs. Nous avons donc souhaité garder des traces des impressions des élèves sur cette caractéristique de la notion bayésienne, non seulement à travers le questionnaire mais après, lors de sa mise en commun, où nous avons enregistré les échanges avec nos dictaphones. En fait, et en prenant compte des séances précédentes, nous nous attendions obtenir les informations les plus riche pas tant des questionnaires, mais des échanges lors de l'étape de bilan. Malheureusement, cette séance s'est finie sans le bilan général. Nous avons dû nous contenter des informations tirées des questionnaires.

La question suivante demande si une couleur fait toujours changer la valeur de la probabilité d'une composition. Pour le cas d'une réponse négative l'élève devait a) repérer le phénomène sur la feuille en indiquant file et colonne et b) l'interpréter.

En amont du cas banal où une hypothèse devient impossible, le retournement de la bouteille n'a pad fait changer des probabilités à six reprises. Ce phénomène, bien que plus simple à repérer sur les graphiques par un trait de ligne horizontal, peut se faire aussi sur le tableau de valeurs des probabilités.

Les réponses des élèves sont partagées (**Tableau 70**). Le grand nombre de réponses incorrectes nous suggère que la question a été mal posée. Toutefois, quelques élèves (quatre sur huit) ayant répondu « oui », ont repéré correctement des itérations où une probabilité restait constante. Ceux ayant répondu « non », ont aussi indiqué le nombre d'itérations correspondant. Pour ce qui concerne les justifications (trois sur quatorze), sauf un élève qui décrit un cas banal, les autres nous sont impossibles d'interpréter.

	La couleur des billes a-t-il toujours modifié les probabilités des compositions ?
Oui	8
Non	3
Sans réponse	3

Tableau 70

La question suivante, numéro six du questionnaire, cherche à compléter la première où nous avons demandé d'estimer la composition de la bouteille. Cette fois-ci la question demande de ranger les compositions selon le degré de croyance, plus l'élève est sûr d'une composition, plus il doit la placer à droite . Cette question est accompagnée d'une demande de justification. Nous cherchons à savoir si le critère de rangement est celui des probabilités ou si par contre les élèves comparent les distances entre l'échantillon et les différentes proportions dans les compositions.

Dans les quatorze questionnaires, huit élèves ont répondu à la question du rangement des compositions, toutes les réponses étant correctes (CDBEA), seules quatre sont accompagnées de justifications. D'après ces justifications, le critère de rangement serait basé dans tous les cas sur les probabilités des compositions et non pas sur la distance des compositions à l'échantillon.

Cette question sur le rangement répond à une inquiétude d'ordre didactique. Lors de la conception de nos situations-problèmes nous avons cherché à placer la probabilité comme un outil de résolution du problème. Plus précisément, un outil décisionnel. Pour ce problème en particulier l'évaluation de la probabilité devient nécessaire pour appliquer le critère décisionnel de probabilité maximale. Ce critère est, à cause de la complexité de la formule de Bayes, beaucoup moins économique que celui de la comparaison des distances des proportions de l'échantillon à celles des compositions.

Pour ce dernier critère décisionnel le concept de probabilité n'est pas indispensable, il suffit de comparer les proportions. Nous avions toujours eu des doutes par rapport au critère choisi pour les élèves pour leurs décisions. Nous ne savions pas si ils se servaient des proportions ou des probabilités.

Les questions suivantes se proposent de traiter le lien entre les deux moyens mentionnés pour répondre au problème, l'un par les probabilités, l'autre par la proximité des proportions. Cet ensemble de questions demande aux élèves d'abord de réorganiser librement les couleurs obtenues lors des retournements et puis d'indiquer si le nouveau rangement ferait changer leur choix de composition. Cette question est suivie d'une demande de justifications.

A cette question a répondu la moitié d'élèves (sept sur quatorze). Les nouveaux rangements effectués par les élèves sont de deux types, soit ils alternent des couleurs O et N (ONONONONOO), soit ils réunissent les retournements de la même couleur (NNNNOOOOOO). Pour les sept élèves, les nouveaux rangements n'impliquaient aucune révision de leurs choix. Les justifications à ces réponses sont variées, parmi les quatre ayant justifié, un élève mentionne la même proportion dans les deux suites, un autre signale que le nouveau rangement ne modifie pas les probabilités finales.

Même si les données sont très faibles en nombre, il semblerait que la proportion dans l'échantillon et les probabilités sont des arguments équivalents pour cette question.

Tel que nous l'avons mentionné auparavant, la séance s'est terminé sans le bilan de l'étape du travail sur l'ordinateur. Dans les deux séances précédentes, les étapes de bilan ont été riches en réflexions des élèves. Malheureusement pour cette séance, nous nous sommes vus privés des échangés dérivés du questionnaire.

Par la suite, et avant de commencer les conclusions pour cette expérimentation, nous reproduisons quelques fragments d'un échange enregistré à la fin de la séance entre un élève et l'enseignant où nous pouvons constater chez cet élèves une persistance déterministe (min 110) :

> **« Que çà dépend totalement de ce que vous avez tiré pour moi çà pourrait bien être quelque chose d'autre c'est pour cela que je n'arrive pas à me prononcer là », « Mais même les élections c'est pas aléatoire », « Même dans la science c'est pas du tout aléatoire rien n'est aléatoire dans tout çà il n'y a rien d'aléatoire rien d'aléatoire »**

Conclusions

Sur la notion bayésienne de la probabilité

La situation-problème de la bouteille a été conçue pour sensibiliser sur la notion bayésienne de la probabilité. Pour son évaluation numérique, nous avons choisi le principe de raison insuffisante. Pour savoir si les élèves ont mobilisé l'interprétation bayésienne de la probabilité, nous avons analysé leurs échanges à l'aide des éléments caractéristiques décrits dans le Chapitre I.

Un des éléments caractéristiques le plus éclairant est la nature de l'objet sur lequel la classe a probabilisé. Les interventions des élèves comme de l'enseignant indiquent que l'objet était bien une hypothèse. Dans les échanges, la classe s'est référée au contenu de la bouteille présente dans la salle. Les discussions étaient donc de caractère épistémique. Il s'agissait de dévoiler le contenu de la bouteille que l'enseignant manipulait devant la classe.

Dans le décryptage de nos enregistrements nous n'avons trouvé aucune trace d'intérêt pour la reproductibilité de l'épreuve, celle-ci étant essentielle à la notion fréquentiste de la probabilité. En d'autres termes, ni l'enseignant ni les élèves se sont intéressés à ce qui arriverait si l'on reproduisait l'expérience. Nous pourrions conclure que la classe a bien employé la notion bayésienne de la probabilité pour se référer au contenu de la bouteille.

Sur la dualité de la probabilité

Toutefois, la notion bayésienne n'a pas été la seule interprétation présente dans cette séance. Lorsque l'on s'intéresse à la couleur d'une bille, sa probabilité peut être du type tant bayésien que fréquentiste, selon l'objet soit un lancer ou la tendance d'apparition d'une couleur donnée. Plus précisément, lorsque la classe s'intéressait à la couleur de la bille pour un retournement (le deuxième par exemple), la probabilité était bayésienne. Elle signifiait une mesure de ce qu'on attend, pour ce lancer, l'apparition d'une couleur donnée. Cette attente, variable, était fonction de ce qu'on connaissait sur la composition de la bouteille.

La probabilité fréquentiste est en principe fixe, caractéristique qui permet de la designer comme étant "objective". Cette probabilité apparaît lorsque la classe s'intéressait à la fréquence d'apparition d'une couleur donnée. Pour l'évaluer, *a priori*, il fallait connaître la composition de la bouteille, cas contraire, on ne pourrait que l'estimer par un grand nombre de retournements.

La dualité de la probabilité est donc présente dans ce problème, l'interprétation bayésienne correspond à la bouteille, la fréquentiste pouvant éventuellement se présenter pour les billes. Pour cette situation-problème nous avions choisi d'institutionnaliser le versant bayésien, bien qu'il serait tout à fait possible d'exploiter sa dualité aussi.

En fait, l'étape d'institutionnalisation dans cette séance ne s'est produite que de manière partielle, rappelons que l'expérimentation a fini lorsque les élèves étaient en train de remplir le questionnaire. Le bilan final prévu, où l'enseignant formalise la notion bayésienne, son évaluation et son évolution ne se sont pas produits. Il reste donc inexploitée une grande partie du travail effectué pour les élèves. Il manque, en effet, le débat qui suit le questionnaire, débat où nous attendions à récolter des renseignements sur les raisonnements des élèves et sur la gestion de l'enseignant.

Néanmoins, même faute de cette étape finale de bilan, quelques conclusions sont possibles. Il semblerait que la dualité ne pose pas des problèmes aux élèves qui, se servent en toute naturalité de la probabilité comme degré de certitude.

En fait, les plus grandes difficultés pour les élèves sont apparues ailleurs, principalement lorsqu'ils devaient décider sur la base d'arguments indéterministes. Ces difficultés semblent répondre à des raisons épistémologiques et didactiques. Pour ce qui concerne les premières, plusieurs auteurs signalent les spécificités de la statistique dans ses critères de validation.Crombie, 1980; Hacking, 1965. Des critères en statistique qui valident une proposition non pas par sa valeur logique mais par le poids de leur argumentation.

Aux difficultés propres à ce genre de démarche nous devrions ajouter celles provoquées par la confusion que produit chez les élèves l'introduction de ce genre de démarche dans une classe de mathématiques. En d'autres termes, aux difficultés propres à ce genre d'argumentation nous devons considérer celles qui prennent leur origine dans un changement de contrat didactique d'une classe de mathématiques.

Les interventions de l'enseignant validant quelques propositions et argumentations des élèves pourraient s'expliquer par la nécessité de la classe d'une approbation de ces nouvaux genre d'argumentation. De cette manière, la ratification de l'enseignant remplirait le vide de l'impossibilité de sa démonstration.

Sur l'évaluation par le principe de raison insuffisante

Cette expérimentation nous a permis de tester certaines variables didactiques, parmi elles, celle qui cherchait à faire émerger le principe de raison insuffisante comme critère d'évaluation d'une probabilité bayésienne. Deux groupes l'ont utilisé de leur propre initiative pour évaluer la première distribution *a priori*. Lorsque ce principe a été présenté à la classe par ces groupes, il n'a pas trouvé de réfutation, mais il manquait un consensus nécessaire pour son acceptation.

En effet, tel que nous venons de le signaler, l'enseignant a dû intervenir pour le valider. Cette tâche de validation de l'enseignant n'a pas requis d'un grand effort de sa part, il a suffi simplement qu'il reprenne les termes et idées de ces deux groupes pour que tout de suite il soit accepté par les élèves.

Le principe de raison insuffisante est, d'un point de vue didactique, important, il met en évidence :

- d'une part l'existence d'autres moyens que les ensembles de référence pour évaluer une probabilité,
- d'autre part, le caractère épistémique de la probabilité. En abordant des cas où la connaissance sur le système est nulle, la valeur de la probabilité ne peut pas être associée à aucune stabilisation de fréquences.

De cette manière, en neutralisant l'intrication des signifiés, les deux interprétations de la probabilité sont séparées. Par exemple, si nous appliquons le principe de raison insuffisante à la première distribution *a priori*, nous n'avons aucun argument pour interpréter ces probabilités comme la fréquence avec laquelle apparaissent chacune des compositions.

Les objets de type hypothèses où l'évaluation se réalise par ce principe permettent donc de dissocier la dualité de la probabilité, une fois bloquée l'interprétation fréquentiste, la nature épistémique de l'objet devient plus évidente.

D'ailleurs, les deux groupes ayant employé le principe de raison insuffisante sont restés bloqués dans une position déterministe. Leurs distributions ne présentaient que deux valeurs logiques, savoir ou ne pas savoir. La première était représentée par des probabilités nulles, les deuxième par l'équirépartition. Il nous semble que ce blocage pourrait s'expliquer par le signifié particulier des probabilités lors de l'application du principe.

En effet, lorsque on applique le principe de raison insuffisante, les probabilités représentent un état de manque de connaissance. A cette état, par opposition, s'associerait la certitude (P(H)=1 ou P(H)=0). Ces valeurs logiques favoriseraient la mise en place du paradigme déterministe.

Il nous semble que pour qu'un changement de paradigme soit possible, il faudrait que les principes du déterminisme soient remis en cause, cela serait possible si l'on réalise un nombre important de retournements. La nécessité de prise en compte de l'information obtenue lors d'un nombre important de retournements pourrait faire du paradigme déterministe un modèle obsolète. En d'autres termes, la nécessité de prise en compte de l'information provenant des retournements créerait une tension sur le modèle qui ferait basculer les élèves vers l'indéterminisme.

Les deux premiers retournements, où les élèves ont employé des jetons, se sont avérés insuffisants pour créer les tensions nécessaires au changement de paradigme. La variable didactique (nombre de retournements) devrait avoir assumée des valeurs supérieures.

Sur le raisonnement par abduction

Ce genre de raisonnement est largement utilisé dans leur vie courante. Par exemple, a chaque fois qu'un indice nous fait réagir, nous serions en principe en train de raisonner par abduction Peirce, 1932. En effet, notre éventuelle action ne serait pas fondée sur la certitude d'une hypothèse mais sur la forte probabilité donnée par l'indice observé.

Bien qu'une grande partie des actions des élèves soit régie par ce genre de raisonnements, il a fallu l'intervention de l'enseignant pour que la classe, lors de l'étape du premier bilan, puisse avancer sur la voie abductive. Cette intervention a fonctionné comme déclencheur des nouveaux arguments des élèves. La réaction immédiate des élèves à l'intervention de l'enseignant nous suggère que cette démarche se trouvait latente chez eux et qu'elle n'émergeait pas pour des raisons de contrat didactique. En d'autres termes, avant l'intervention de l'enseignant, les élèves pouvaient avoir perçu ce genre d'argument, comme étant non attendus dans une classe de mathématiques. L'intervention de l'enseignant aurait fonctionné comme une approbation.

Dans cette séance, les élèves semblent ne pas avoir eu des difficultés ni à raisonner de manière abductive ni à accepter les arguments abductifs provenant d'autres élèves.

Sur la formule de Bayes

Un des points les plus problématiques pour cette expérimentation a été la formule de Bayes. Elle doit être considérée, nous en sommes convaincus, comme un outil effectuant dans le champ numérique ce que l'abduction fait dans le champ de l'argumentaire. En d'autres termes, un moyen pour quantifier les différentes étapes du raisonnement par abduction.

D'ailleurs, cette formule présente des difficultés à cause de la variété d'éléments qui la composent. Variété qui fait de leur suivi ou contrôle sémantique une tâche fort complexe. Par exemple, la formule de Bayes fait intervenir quatre types de probabilités :

$P_H(couleur)$. Cette probabilité s'évalue en admettant une hypothèse, elle peut être, bayésienne ou fréquentiste.

$P_{couleur}(Hypothèse)$. Cette probabilité porte sur une hypothèse, aussi connue comme probabilité des causes. Par rapport à la précédente, il y a une inversion, on admet un effet et on probabilise sur sa cause. Elle est de nature bayésienne.

$P(couleur)$. Couramment connue comme probabilité totale, elle représente une attente moyenne en fonction du degré de connaissances des possibles hypothèses. Elle est de nature bayésienne.

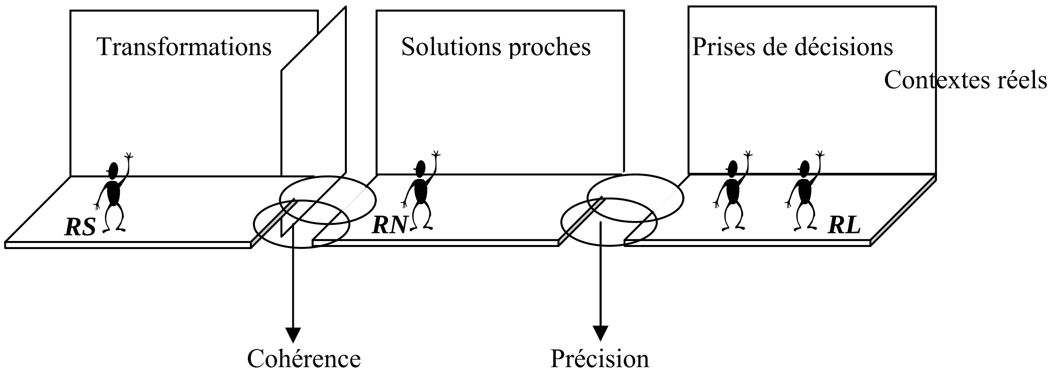
$P(H \cap couleur)$ et $P(couleur \cap H)$. Sémantiquement difficiles à saisir (Gras & Totohasina, 1995), elles représentent la probabilité de l'occurrence des deux phénomènes, elles sont aussi de nature bayésienne.

Il nous semble que la compréhension intégrale de cette formule requiert un travail au préalable sur chacune de ces probabilités. Très probablement, une sensibilisation progressive à travers des situations-problèmes spécialement conçues pour ces différents éléments. Il est très improbable qu'un groupe d'élèves puisse saisir le signifié de la formule de Bayes dans une seule séance, sans jamais avoir été sensibilisé aux signifiés des éléments qui la composent.

Notre situation-problème ne prétend pas en être une pour la construction de cette formule. Pour cela, nous aurions du réfléchir à une suite de situations-problèmes. Un tel projet nous est paru improbable à mettre en place, principalement à cause du nombre de séances nécessaires. Le programme de mathématiques consacre peu d'attention aux aspects sémantiques des concepts de statistique et cela se projette sur le temps consacré en classe. Notre probabilité de trouver un enseignant prêt à céder un nombre non négligeable de séances à des concepts qui ne font pas partie du programme étaient très faible. Nous avions ainsi renoncés à ce projet et nous nous sommes contenté d'expérimenter juste sur quelques aspects de la problématique bayésienne.

En principe, lors de la conception de cette situation-problème, nous avions retenu l'introduction de la formule de Bayes comme une réponse à la nécessité de mathématiser le procédé abductif. Ces idées se trouvent schématisées en la Figure 15 que nous reproduisons ci-contre. Le passage au registre numérique (RN) serait motivé par la nécessité de précision, et celui au registre symbolique (RS) par la nécessité de critères de transformation. De cette manière, la formule de Bayes représenterait en termes symboliques les critères de transformation fournis par le raisonnement par abduction.

Représentation de caractéristiques communes aux deux interprétations de la probabilité



Dans la conception de la situation-problème, cette formule étant ensée être présentée par l'enseignant, de manière assez sommaire, où l'axe de son introduction reposait sur son utilité pour donner de manière précise les valeurs des probabilités après l'arrivée de l'information. Il était envisagé aussi d'observer en classe, et sur la feuille de calcul, la correspondance entre les calculs effectués par l'algorithme et les distributions de jetons.

A cet effet, nous avions ordonné la présentation de la formule. Cette présentation visait à faciliter le repérage de la fonction de transformation des probabilités. Le schéma consistait en présenter là formule en trois blocs. Il y avait: une probabilité avant l'information (*a priori*), un facteur de transformation et puis la probabilité *a posteriori*. Cette présentation suit le schéma de la formule transcrite ci-contre :

$$\boxed{P_O(K)} = \boxed{\frac{P_K(O)}{P(O)}} \times \boxed{P(K)}$$

État final des croyances | Taux de changement des croyances | État initial des croyances

L'emploi de la formule de Bayes est ainsi similaire à celui d'un grand nombre d'utilisateurs de méthodes statistiques qui connaissent les principes de la méthode mais pas les détails de l'algorithme. Un des objectifs pour cette séance était précisément d'observer les possibilités d'une telle introduction de la formule de Bayes.

Tel que nous l'avons mentionné auparavant, l'enseignant a choisi de développer la démonstration de la formule. Cet intérêt de l'enseignant pour justifier la formule de Bayes, nous semble répondre à une habitude du métier, où l'introduction d'un objet devrait être accompagnée, si possible, de sa démonstration. Cette habitude a influencé l'enseignant à consacrer un temps non négligeable de la séance à cette démonstration.

Chapitre IV : Conclusions et perspectives

4.1 Introduction

Dans ce dernier chapitre nous nous intéressons à quelques conclusions tirées de notre recherche et à des possibles perspectives qu'en découlent. La présentation des conclusions suit de manière générale l'organisation de cette recherche, premièrement les épistémologiques, deuxièmement celles concernant l'analyse des manuels et troisièmement une partie consacrée aux expérimentations. Finalement nous proposerons quelques perspectives où nous inclurons des questions non traitées dans notre recherche.

4.2 Enquête épistémologique

Notre sujet de thèse est né de la lecture de deux textes, d'une part l'ouvrage de l'épistémologue Ian Hacking Hacking, 2002 sur l'émergence de la probabilité où l'auteur développe les différentes circonstances culturelles allant mené à la construction de la notion duale de la probabilité ; et d'autre part, un article du chercheur Bruno Lecoutre Lecoutre, 2005 où il décrit quelques erreurs courantes lors de l'interprétation de tests d'hypothèses, principalement à cause de la confusion entre les notions bayésienne et fréquentiste.

De ces lectures, nous avons tiré quelques conclusions :

- Premièrement, si des erreurs sont possibles de repérer, alors l'interprétation de la probabilité ne serait pas une question d'approches personnelles mais une caractéristique de la situation-problème, cas contraire on ne parlerait pas en termes d'erreurs mais de points de vue.
- Deuxièmement, si des telles caractéristiques sont plausibles à préciser, nous pourrions nous en servir pour l'élaboration de situations-problèmes destinées à l'enseignement de la probabilité.

Ces premières réflexions nous ont mené à approfondir l'enquête épistémologique. C'est ainsi que, motivés pour la recherche de caractéristiques propres à chaque interprétation de la probabilité, nous avons fait une incursion dans le domaine de l'inférence, tant classique comme bayésienne. Dans cette enquête nous avons trouvés des positions variées par rapport aux deux interprétations de la probabilité, allant des radicales (Jaynes, Popper, etc.) aux éclectiques (Hacking, Shafer, etc.), les premiers n'acceptant qu'une seule des écoles, les deuxièmes admettant la cohabitation des deux approches.

Dualité incontournable

Au même temps que nous apprenions sur les différences entre l'approche fréquentiste et bayésienne, l'hypothèse d'une dualité incontournable se consolidait. Les caractéristiques mêmes du concept nous ont mené à postuler qu'en amont de toute intention pédagogique, les deux interprétations de la probabilité se présenteraient en classe. Une explication à ce phénomène se trouverait en ce que nous avons dénommé l'intrication des signifiés.

L'intrication des signifiés se manifeste lors de l'évaluation d'une probabilité bayésienne. Elle consisté à évaluer la certitude d'une proposition par la fréquence de son apparition. Un exemple simple : Parmi deux routes possibles, nous souhaitons choisir la moins dangereuse. Pour l'une, le taux d'accidents mortels étant de deux voitures sur dix mille qui l'empruntent, pour l'autre, de six sur dix mille. Si nous nous intéressons à évaluer notre probabilité d'un accident mortel sur chacune de ces routes, la fréquence d'accidents devient une raison pour quantifier que l'une et moins dangereuse que l'autre. Mais dans aucun cas, cette probabilité ne signifie pas que nous allons mourir avec une moyenne de deux sur dix mille ou de six sur dix mille.

La structure de ce genre de situations serait : on s'intéresse à la réalisation d'un phénomène et n'ayant pas d'autre information que la fréquence de son apparition, on utilise cette dernière pour évaluer combien on croit en sa réalisation. Ce procédé où la fréquence devient raison à croire nous l'avons dénommé référenciation par le principe fréquentiste.

L'intrication de signifiés n'est pas la seule raison pour postuler l'inévitabilité de la dualité de la probabilité en classe. D'autres situations sont possibles pour que les deux notions soient présentes, une des plus courantes consiste en la proportion de cas favorables et possibles. En effet, cette proportion devient critère d'évaluation tant pour la notion fréquentiste comme pour la bayésienne, voilà donc un autre point d'intersection des deux notions de la probabilité.

De plus, la dualité d'interprétation se trouve déjà dans plusieurs langues occidentales (anglais, espagnol et français, par exemple). Dès l'origine même du concept jusqu'à nos jours, le terme probabilité fut utilisé pour se référer tant à la fréquence d'apparition d'un phénomène comme à un degré de certitude. Il semblerait que les caractéristiques culturelles de l'époque de son émergence ont fait de cette notion un concept à deux visages impossible à dissocier. Cette hypothèse Hacking, 2002 et l'intrication de signifiés citée, nous ont mené à postuler l'inévitable reproduction de cette dualité en classe.

Pour confirmer notre hypothèse d'incontournable dualité de la probabilité, il nous manquait des éléments nous permettant d'identifier chacune des notions de la probabilité, c'est ainsi que nous nous sommes consacrés à chercher des éléments caractéristiques à chacune des interprétations de la probabilité.

Les éléments caractéristiques

La première étape de notre recherche a consisté donc à repérer des caractéristiques propres à chacune des interprétations de la probabilité. Derrière cette recherche se cachait une hypothèse, celle de l'objectivité d'approche par les caractéristiques du problème, hypothèse que nous avons tiré lors de nos premières lectures. De cette manière, si une identification était possible, elle le serait par les caractéristiques de la situation-problème. Cette recherche de caractéristiques différenciatrices nous a permis d'identifier quatre éléments, qu'en se complémentant, nous permettent d'identifier l'interprétation de la probabilité associée au contexte d'un problème.

A l'intention d'identification d'éléments caractéristiques, nous avons joint celle de repérage de la présence de la dualité de la probabilité au long de son histoire. Pour articuler ces deux objectifs (présentation d'éléments caractéristiques et repérage historique) nous nous sommes inclinés pour une présentation du type chronologique, où nous reprendrions quelques extraits d'auteurs de référence au même temps que nous présenterions nos éléments caractéristiques.

Par sa signification, le premier élément en importance reste le type d'objet sur lequel porte la probabilité, nous l'avons décliné en trois catégories : série infinie, épreuve générique et hypothèse. Pour caractériser la série infinie, nous reproduisons la définition de Richard von Mises. D'auprès lui, la probabilité fréquentiste est une caractéristique d'une série infinie satisfaisant deux axiomes :

- Axiome de convergence : La fréquence relative du caractère (ou propriété) observé dans le *collective* tends vers l'infini.
- Axiome de randomisation : La limite de convergence ne se voit pas affectée par une sélection d'un sous ensemble infini quelconque à la condition que la règle de sélection soit fixe.

Pour ce qui concerne la notion bayésienne, deux types d'objets sont possibles, l'un l'épreuve générique, l'autre l'hypothèse. Ces deux catégories se différencient principalement en que, pour la première, l'observateur dispose de moyens pour se représenter la reproduction de l'épreuve, delà, son caractère générique. Le cas dénommé hypothèse se caractérise pour un contexte ne disposant d'aucun ensemble sur lequel se référencier, quelques auteurs le dénomment cas unique Hacking & Dufour, 2004 .

Le deuxième élément retenu concerne la valeur logique associée, deux cas sont possibles, pour la notion fréquentiste les valeurs étant dichotomiques (vrai ou faux), pour la notion bayésienne, l'intervalle continu entre le vrai et le faux. En effet, pour la notion fréquentiste, chaque reproduction de l'événement est, en termes logiques, une fonction à deux valeurs possibles, soit l'événement se produit (vrai) soit il ne se produit pas (faux). De plus, la valeur de la proportion d'occurrences en la série est unique, elle est donc aussi évaluable en termes dichotomiques. La notion bayésienne est dans ce sens plus large, en tant que mesure de certitude, la valeur de la probabilité non seulement peut prendre des valeurs de certitude totale (vrai-faux), mais aussi partielle ((0,1)). De cette manière, plus sûr on est d'une proposition, plus proche de 1 sera la probabilité.

Le troisième élément concerne le type de raisonnement mobilisé, les deux notions se différencient sur ce plan aussi. Pour ce qui concerne la notion fréquentiste, l'évaluation de la proportion de l'occurrence de l'événement au but d'un nombre infini de répétitions de l'expérimentation est un procédé déductif, en admettant un ensemble d'hypothèses, on déduit la valeur limite de la convergence. Pour la notion bayésienne deux cas sont possibles. Pour une épreuve générique, l'évaluation se réalise par référenciation, qui à son tour décline en deux possibilités, soit la référence est un ensemble fini, soit elle est infinie. Le premier type de référence, typique aux axiomes de Kolmogorov, prend en compte la proportion de cas favorables sur cas possibles pour évaluer la probabilité. Le deuxième type, la référenciation par des ensembles infinis, se produit lorsque l'on évalue la probabilité par le principe fréquentiste. L'autre cas pour la notion bayésienne le constitue l'hypothèse, nous y avons réuni des situations inférentielles, où l'information constatée mène à une révision des degrés de certitude portés sur une hypothèse. Le genre de raisonnement associé est dénommé comme abduction.

Épreuve générique et hypothèses sont étroitement liées. La principale différence réside en que pour la première, le contexte du problème permet la représentation du caractère générique de l'épreuve, de cette manière, elle est considérée comme une épreuve parmi d'autres. Cette inclusion de l'épreuve dans un ensemble plus vaste (fini ou infini) permet de l'évaluer par la proportion qu'elle représente dans l'ensemble. L'objet de type hypothèse par contre, par manque d'un ensemble de référence, empêche une représentation de type générique. L'évaluation de sa probabilité s'effectue par d'autres sources d'information, écarté son caractère générique, on la quantifie par ce que l'on connaît d'elle. De cette manière, apparaissent des critères plus personnels, tels que le principe de raison insuffisante, traité dans nos expérimentations.

Une épreuve générique peut dériver en une hypothèse. Par exemple, en la situation-problème de la bouteille, si l'enseignant avait donné aux élèves la composition de l'urne, ils auraient évalué la première distribution de probabilités par référenciation. Puis, lors des retournements, la nature de l'objet tournerait vers une hypothèse. L'épreuve perdrait son caractère générique et l'information renseignerait sur la bouteille en particulière. La révision ou réévaluation des probabilités se réaliserait non plus par référenciation mais par abduction.

Pour l'évaluation des hypothèses nous nous sommes servis du principe de raison insuffisante, mais il n'est pas le seul critère possible. Par exemple, lors de l'expérimentation de la situation de la bouteille, si l'enseignant avait donné l'impression de fouiller dans l'urne en cherchant à obtenir une proportion particulière, les élèves auraient probablement privilégié une composition en détriment d'autres, et cela en fonction de ce qu'ils connaissent de leur enseignant, argument tout à fait légitime sous l'approche bayésienne.

Ceci constitue un exemple de la subjectivité des critères d'évaluation d'une probabilité bayésienne. Ces critères pourraient paraître arbitraires, néanmoins, ils ne sont que des raisons pour quantifier une mesure de certitude, de la même manière que le fait la proportion de cas favorables et possibles pour évaluer une épreuve générique.

Dans ce sens, il nous semble important de rappeler que pour l'approche bayésienne, la première évaluation de la probabilité n'est pas aussi efficace que de compter avec un critère de révision permettant de réévaluer les probabilités en fonction de ce que l'on apprend sur l'objet (formule de Bayes).

La quatrième différence concerne les principes utilisés pour l'évaluation d'une probabilité. Pour la probabilité fréquentiste, et en excluant les cas où la proportion est estimée, la probabilité s'évalue par le principe d'équiprobabilité, qui au même temps renvoie à l'utilisation de la formule de Laplace. Pour la probabilité bayésienne, nous nous sommes concentrés sur trois principes. Deux pour l'épreuve générique et un pour l'hypothèse. L'épreuve générique pouvant être évaluée par le principe d'équipossibilité si la référence est un ensemble fini, ou par le principe fréquentiste si la référence est un ensemble infini. Pour ce qui concerne l'hypothèse, les sources sont diverses, nous nous sommes restreints à l'utilisation du principe de raison insuffisante, le **Tableau 71** résume les quatre différences repérées.

Quatre différences entre les interprétations de la probabilité

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="5">Interprétations</td></tr>
<tr><td colspan="2">Fréquentiste</td><td colspan="3">Bayésienne</td></tr>
<tr><td rowspan="4">Diff éren ces</td><td>Objet sur le quel probabiliser</td><td colspan="2">Série infinie</td><td colspan="2">Epreuve générique</td><td>Hypothèse</td></tr>
<tr><td>Valeurs logiques portées sur ces objets</td><td colspan="2">Vrai- faux</td><td colspan="3">Intervalle continue [0,1]</td></tr>
<tr><td>Type de raisonnement</td><td colspan="2">Déduction</td><td colspan="2">Référenciation</td><td>Abduction</td></tr>
<tr><td>Quelques principes de quantification</td><td>Hypothèses du modèle</td><td colspan="2">Equipossibilité (Ensemble fini)</td><td>Fréquentiste (ensemble infini)</td><td>Raison insuffisante</td></tr>
</table>

Tableau 71

D'ailleurs, la différenciation des deux interprétations de la probabilité déclinée en éléments caractéristiques est associée à notre intérêt pour une approche statistique de la probabilité où l'interprétation d'un calcul est fondamentale. En d'autres termes, la probabilité n'est pas seulement un ensemble d'axiomes permettant d'effectuer des calculs mais un concept à deux dimensions, l'une sémantique et l'autre calculatoire. Les éléments caractéristiques décrivent les liens entre ces deux dimensions. Par exemple, la dimension sémantique est précisée par l'élément caractéristique *nature de l'objet*, la dimension calculatoire par les *principes d'évaluation*.

Dans cette approche interprétative de la probabilité, nous identifions une relation hiérarchique entre ces deux dimensions où la sémantique se situe par-dessus de la calculatoire, c'est la première qui définie le champ du possible pour la deuxième. En d'autres termes, c'est l'approche interprétative qui définie le genre de situation à traiter et les possibles moyens pour évaluer la probabilité.

Une autre observation concerne la formule de Laplace. Pour certains, la probabilité admet trois acceptions, fréquentiste, bayésienne et laplacienne. Pour ce qui concerne la troisième définition, l'approche assumée dans notre recherche la situe comme un algorithme de calcul, un outil permettant le passage au nombre de la probabilité et non pas une définition en termes sémantiques. Cette formule est, dès notre approche, d'intérêt lors de l'évaluation d'une probabilité, mais dans aucun cas elle ne serait porteuse de sens. De cette manière, dès une approche interprétative, la probabilité reste dans une dualité de signifié, fréquentiste et bayésienne.

Repérage de la dualité

Le deuxième axe de notre enquête épistémologique a consisté à repérer la dualité au long de l'histoire de la probabilité. Pour cette enquête nous avons procédé en deux étapes, la première a consisté en une lecture épistémologique nous ayant servi de guide pour la deuxième étape, la lecture de textes d'auteurs de référence.

Dans la première étape, nous reconnaissons l'influence de plusieurs auteurs, parmi eux : Ian Hacking, Glen Shafer, Andrew Dale, Miche Armatte, Alistair Crombie, Alain Desrosières, Claudine Tiercelin et M. A. Todhunter. Sur cette base épistemologique, nous avons abordé la lecture de textes historiques et contemporains. De cette deuxième étape, nous avons retenu quelques textes afin de montrer la présence des deux interprétations de la probabilité dès débuts même de cette notion.

Le **Tableau 72** résume les textes retenus par auteur, année de publication, titre et sujet. Les ouvrages les plus anciens, nous les avons trouvés sur Internet, dans quelques bibliothèques numérisées. Les autres, nous les avons consultés sur de support papier. Du à sa gratuité et commodité, nous avons privilégié le support numérisé, cela explique que quelques ouvrages sont en anglais.

Sujet développé par auteur

Auteur	Année	Titre	Sujet
B. Pascal	*1670*	*Pensées*	*Notions bayésienne et fréquentiste*
Nicole & Arnaud	*1662*	*La logique de Port Royal*	*Evaluation. Ensembles de référence*
G. Leibniz	*1765*	*New essays concerning human understanding*	*Notion bayésienne*
J. Bernoulli	*1713*	*The art of Conjecturing*	*Notion fréquentiste*
T. Bayes	*1763*	*An essay towards solving a problem in the doctrine of chances*	*Notion bayésienne*
P. Laplace	*1795*	*Essai philosophique sur les probabilités*	*Notions bayésienne et fréquentiste*
J. Keynes	*1921*	*A treatise on probability*	*Notion bayésienne*
B. de Finetti	*1937*	*La prévision: ses lois logiques, ses sources subjectives*	*Notion bayésienne*
R. von Mises	*1928*	*Probability, statistics and truth*	*Notion fréquentiste*
K. Popper	*1959*	*The logic of Scientific Discovery & The propensity interpretation of probability*	*Notion fréquentiste*

Tableau 72

Nous n'allons pas reproduire ici les citations réalisées dans le premier chapitre où nous avons transcrit quelques paragraphes montrant l'utilisation d'une ou de l'autre approche dans chacun de ces ouvrages. Nous nous contenterons d'affirmer que l'enquête réalisée sur les deux types de sources (épistémologiques et historiques) nous confirment la constitution duale de la probabilité depuis son émergence à nos jours.

Cette enquête épistémologique a été fondamentale à notre sujet de recherche. Les éléments caractéristiques d'y repérés, nous ont permis de nous outiller non seulement pour bien différencier chacune des approches mais aussi pour aborder la construction de situations-problèmes.

De cette manière, les éléments caractéristiques sont devenus une guide à l'élaboration de situations-problèmes, en satisfassent les conditions données par ces éléments nous avons choisi des problèmes accords à chaque approche. De plus, et en s'agissant de caractéristiques relativement objectives à repérer, elles nous ont permis d'analyser l'interprétation associée à des situations-problèmes déjà construites, telles que les exercices des manuels. Cela a été le sujet du deuxième chapitre, où nous avons analysé la présence de la dualité de la probabilité dans quatre manuels de lycée français. C'est en nous basant sur ces éléments que nous confirmons la présence des deux interprétations dans les manuels de lycée français.

La probabilité, une aide à la décision

Une autre conclusion de l'enquête épistémologique concerne le rôle de la probabilité comme outil d'aide à la décision. Dans ce sens, les premières textes consultés ont été éclairant (Bayes, 1763; Leibniz, 1765; Nicole & Arnaud, 1662; Pascal, 1670), les traces d'un intérêt pour rationaliser l'incertain par la probabilité sont récurrents dans ces auteurs et même dans des textes contemporains (Gärdenfors et al., 1988). La probabilité donc, se place comme un outil pour agir rationnellement devant l'incertain, cette fonction épistémologique de la probabilité nous l'avons transposée au didactique, de cette manière et dès une approche élargie de la dialectique outil-objet (Douady, 1986) nous avons placé la probabilité dans nos situations-problèmes comme un outil décisionnel.

Ce rôle de la probabilité d'aide à la décision renforce l'importance des interprétations de la probabilité. En effet, une prise de décision requiert que la probabilité soit habillée d'un signifié, sans son interprétation, l'argument décisionnel devient incomplet. Dans ce sens, nous avons parlé d'une relation de nécessité entre une prise de décision et l'interprétation associée à un problème, sans la deuxième, on ne peut développer la première.

De cette relation de nécessité nous avons tiré profit dès un point de vue didactique, pour justifier leurs choix, les élèves ont été obligés d'expliciter le sens attribué à la probabilité. C'est donc, en demandant des prises de décisions dans les situations-problèmes que nous avons trouvé un moyen didactique pour arriver aux interprétations de la probabilité.

4.3 Analyse de manuels

L'analyse de manuels eut plusieurs objectifs, premièrement confirmer la présence de la dualité de la probabilité dans les exercices des manuels, deuxièmement, comprendre comment les manuels résoudraient le conflit entre dualité incontournable et choix d'interprétation du programme. Troisièmement, nous approcher au genre de situations auxquelles les élèves de BTS auraient été confrontés lors de leur passage pour le lycée. Cette approximation par les manuels nous renseignerait non seulement de possibles connaissances disponibles chez les élèves, mais aussi des profils d'exercices habituels, d'attentes réciproques entre enseignant et élèves, etc. Enfin, en analysant les manuels, nous nous sommes outillés pour nos expérimentations.

L'analyse de manuels a été précédée par une brève description de l'approche proposée par les rédacteurs du programme. Pour cela, nous avons transcrit, quelques paragraphes des programmes et des documents d'accompagnement, documents tous datant de l'année 2001 et en vigueur l'année de notre expérimentation (2005-2006). Le projet explicité dans ces documents consistant en associer le calcul de la probabilité à l'interprétation fréquentiste de la probabilité.

La dualité dans les manuels

Pour analyser la présence des deux interprétations de la probabilité dans les manuels, nous avons considéré un ensemble de variables regroupées en deux catégories. La première de ces catégories nous renseignant de caractéristiques générales des exercices (taille, nombre de questions, type de représentation proposées, etc.), la deuxième, nous permettant d'identifier la notion de la probabilité sous-jacente au problème (nature de l'objet, manipulation d'objets, hypothèse du modèle, etc.). Cette deuxième catégorie a inclus d'autres variables nous renseignant sur la manière de résoudre le conflit entre une dualité incontournable et les directives des programmes (décision, interprétation de la probabilité, calcul de P, etc.).

Pour cette analyse nous avons relevé de manière exhaustive les exercices du chapitre probabilité de quatre manuels de la première année de lycée de l'année 2001 (filières ES et S). Le choix des manuels a été arbitraire, nous ne nous sommes basé ni sur la proportion de ventes de chaque groupe éditorial ni d'aucun autre moyen nous permettant d'estimer la représentativité de ces manuels dans l'ensemble. Notre objectif n'était pas de valider nos inférences par des méthodes quantitatives, mais plutôt d'explorer la place des interprétations de la probabilité dans les manuels français.

Toutefois, s'une inférence est-il possible, elle pourrait s'expliquer par une hypothèse d'homogénéité entre manuels. En effet, nous pourrions supposer une certaine similitude entre manuels de la même filière due à la nécessité des manuels de répondre aux normes officielles. Toutefois, l'hypothèse d'homogénéité a été partiellement réfutée, parmi les quatre manuels retenus (deux pour la filière ES et d'autres autres pour la filière S) nous avons trouvé un manuel (Bréal ES) don ses exercices se différencient significativement de ceux des autres trois manuels.

La méthode utilisée pour nos analyses a été l'analyse statistique implicative (Gras et al., 1996; Gras & Bailleul, 2000). Pour le traitement de données nous nous sommes servis du logiciel CHIC développé par l'équipe de Régis Gras (Gras et al., 2005). La méthode comme le logiciel sont amplement diffusés en didactique des mathématiques en France, pour plus d'information sur la méthode et le logiciel, nous renvoyons au lecteur à la bibliographie citée.

Notre sujet s'ajustait à genre de problèmes traités par la méthode A.S.I. Nous disposions de données dichotomiques et d'hypothèses de structure implicative. Par exemple, pour ce qui concerne les données, nous nous sommes intéressés à savoir si, dans un exercice, la probabilité calculée était accompagnée d'une demande d'interprétation. La réponse admettait deux possibilités (non : si l'interprétation n'était nécessaire pour répondre à aucune question de l'exercice, oui : cas contraire). Nous avons structuré les questions de manière telle à que les modalités constituent une partition où la réponse à chaque modalité admettait toujours une réponse binaire (oui, non).

Pour ce qui concerne les hypothèses, en effet, elles répondaient à une structure implicative. Nous avions d'une part, une hypothèse principale à confirmer, nous trouverions d'exercices dans manuels correspondants aux deux approches, bayésienne et fréquentiste. D'autre part, les caractéristiques de chaque type d'exercices nous renseigneraient de la manière de surmonter la contradiction entre la dualité incontournable et l'approche du programme.

En effet, un pourcentage non négligeable d'exercices de type épreuve générique ou hypothèse nous aurait confirmeé notre hypothèse sur le caractère incontournable de la dualité. Néanmoins, cette observation univariée ne nous renseignerait pas sur la manière de résoudre le conflit entre ce constat et les directives des programmes. C'est ainsi que nous avons considéré quelques hypothèses implicatives, où la confirmation s'effectuerait de manière partielle. Par exemple, une implication *a priori* fut « si l'exercice est de nature bayésienne, alors l'attention se centrera sur les calculs et non pas sur l'interprétation associée ». Cette implication pourrait devenir impossible à valider sur la totalité des exercices, nous avions donc trouvé comme alternative une validation partielle, par une quasi implication.

Quelques quasi implications, ont été proposées à l'avance, d'autres découvertes lors du traitement des données. Pour ce qui concerne les premières, le manque d'analyse au préalable nous a empêché de quantifier *a priori* la force de ces règles, en termes de seuil attendu, nous les avons proposé sans le seuil correspondant.

Pour nos analyses nous nous sommes servis des trois méthodes disponibles dans le logiciel, arbre de similarités, graphe implicative et arbre cohésitif. Les résultats obtenus se confirment et complémentent par ces trois méthodes. Nous reproduisons quelques conclusions de ces analyses (voir Chapitre II) :

L'hypothèse de dualité incontournable se confirme sur ces quatre manuels. Trois groupes d'exercices ont été repérés. Un groupe appelé fréquentiste, un deuxième bayésien et un dernier sans contexte.

Le groupe fréquentiste n'étant pas constitué exclusivement d'exercices de nature fréquentiste, il contient des exercices ambigus, où la nature de l'objet est dans une partie de l'énoncé fréquentiste, dans l'autre, une épreuve générique. Pour ces derniers cas, la fréquence d'apparition est utilisée pour évaluer la probabilité de l'épreuve générique (référenciation par des ensembles infinis : principe fréquentiste).

Le groupe bayésien se caractérise par des objets de nature épreuve générique. L'évaluation de la probabilité est couramment effectuée sur des ensembles finis. Ces exercices se concentrent sur des tâches calculatoires, ensemblistes ou liées aux axiomes de Kolmogorov. L'interprétation de la valeur de la probabilité demandée à l'élève n'est pas adaptée aux questions posées dans l'énoncé.

Un dernier groupe a été repéré dans les manuels, il s'agit d'exercices dépourvus de contexte. Pour ces exercices il est impossible de déterminer avec précision l'interprétation de la probabilité associée. En effet, les éléments caractéristiques, principalement la nature de l'objet sur lequel porte la probabilité, sont des éléments du contexte de l'exercice. De cette manière, un exercice manquant de contexte reste inaccessible à notre analyse. Ces exercices semblent avoir pour objet, l'entraînement ou le renforcement de tâches calculatoires.

Le schéma de trois types d'exercices (fréquentiste, bayésien et sans contexte) est représentatif des trois des quatre manuels analysés, le dernier s'est avéré quelque part original devant nos variables. Ce manuel (Bréal ES) se différencie des autres sur plusieurs plans et même nos variables semblent parfois inappropriées pour caractériser ses exercices. Toutefois, et malgré la singularité de ses exercices, nos analyses nous ont permis de repérer deux des trois groupes d'exercices (fréquentiste et bayésien) dans ce manuel.

D'autres caractéristiques des exercices

L'analyse de manuels nous a renseigné sur les exercices auxquels étaient confrontés les élèves du BTS avec qui nous avons réalisé nos expérimentations lors de leur passage par le lycée. Même si relativement nombreux, les exercices du chapitre probabilité ont tendance à se ressembler.

- Sur le genre de situations. Les exercices proposent des situations basées sur des jeux. Cartes, dés et billes, font partie des objets les plus évoqués dans les énoncés. Rares sont les exercices où le concept de la probabilité est appliqué à une situation de la vie quotidienne des élèves ou au monde du travail.
- Sur les décisions. Les exercices se caractérisent par une approche descriptive devant l'incertitude. Les énoncés ne proposent pas de prendre de décisions avec les probabilités calculées. Le schéma pouvant s'esquisser comme suit, l'élève est devant un contexte d'incertitude, on lui demande d'effectuer des calculs probabilistes et abruptement l'exercice fini après les calculs. Même dans les exercices basés sur les jeux, la tendance est de ne pas aborder des stratégies plus convenables que d'autres pour gagner le jeu. En particulier, l'espérance mathématique subit une transposition importante, elle passe d'un outil décisionnel dans les textes épistémologiques consultés à un objet d'intérêt calculatoire dans les manuels, où sa principale utilité consiste à effectuer des calculs sur certaines valeurs, par exemple calculer la valeur de la variable aléatoire pour que le jeu soit équitable.
- Sur l'indéterminisme. Bien que nos variables n'ont pas été conçues pour repérer le paradigme métaphysique associé à chaque exercice, les analyses *a priori* nous suggèrent une approche particulière pour la probabilité. Dans les exercices de ces manuels, principalement en trois d'entre eux, tout se passe comme si il n'y aurait pas d'incertitudes autour de la probabilité. Par exemple, les exercices sont organisés pour que, lors de l'évaluation d'une épreuve générique, un seul ensemble de référence soit évoqué. De cette manière la valeur de la probabilité est univoquement déterminée. De plus, l'existence d'un seul ensemble de référence fait que le type de raisonnement est une déduction plutôt qu'une référenciation. A ces caractéristiques se joint l'absence de prises de décisions et l'évocation des contextes à la place d'une manipulation réelle. Des éléments que, agissant ensemble, semblent faciliter le détournement de questions indéterministes vers une approche "déterministe" de la probabilité.

L'analyse des manuels nous a permis de constater la présence de la dualité de la probabilité dans les exercices du chapitre probabilité des manuels de la première année de lycée. De plus, elle nous a renseigné du genre d'exercices auxquels, enseignant et élèves, étaient habitués à travailler en classe. Nous admettons que la classe de BTS où nous avons effectuées nos expérimentations s'est vue confrontée à ce genre d'exercices lors de son passage pour le lycée. De cette manière, l'information tirée des analyses des manuels nous a été d'utilité pour nos expérimentations.

4.4 Expérimentations en BTS

L'étape d'expérimentation s'est avérée complexe, en particulier celle de trouver un enseignant près à céder quelques séances pour expérimenter sur notre sujet. En principe, nous avions choisi d'expérimenter en lycée, mais après quelques tentatives infructueuses nous nous sommes décidés pour une formation ayant des contraintes moins restrictives.

En effet, nous avions trouvé deux enseignants dans un lycée des alentours de Paris. Lors d'un entretien, ces deux enseignants se étaient manifesté favorables à réaliser des expérimentations autour de l'interprétation de la probabilité dans leurs classes de la filière scientifique, l'un dans une classe de Première, l'autre de Terminal. Afin d'analyser les possibilités d'expérimenter nos situations-problèmes, nous avons procédé à effectuer deux pré-expérimentations, une pour chaque classe. Pour la classe de Première, nous avons expérimenté la situation de la bouteille, mais à partir de l'approche de Brousseau, pour la classe de Terminale, nous avions proposé un jeu, où les élèves mettaient en place des notions élémentaires d'un test χ^2.

Ces expérimentations, dont nous conservons les enregistrements des séances, feront l'objet d'une analyse détaillée dans le futur. Ces deux séances et quelques entretiens, même très riches, nous ont conduit à conclure que nos expérimentations seraient fortement conditionnées par les contraintes qui s'exercent sur cette filière de lycée. Nous avons donc cherché d'autres enseignants intéressés, mais disposant de marges de manœuvre plus larges sur les sujets du programme.

Après plusieurs rencontres, nous avons fait connaissance de cet enseignant d'une formation de BTS en électrotechnique, qui s'est manifesté très motivé dès le début pour tester des situations-problèmes autour de la probabilité. L'établissement où cet enseignant[20] exerçait ses fonctions se trouve en province. Nous avons donc résigné d'effectuer des pré expérimentations (au moment de nos expérimentations, nous habitions à Paris).

Néanmoins, les échanges ayant précédé nos expérimentations ont été nombreux, bien qu'en grand nombre non enregistrés. Entre la première rencontre (mai 2006) et la première

[20] Cet enseignant est diplômé d'un DEA en statistique.

expérimentation (Novembre 2006), nous avons discuté d'abord sur des questions épistémologiques autour de la dualité de la probabilité et puis, dans un deuxième temps, sur chacune des situations-problèmes. Les caractéristiques particulières de ce projet nous ont mené à non seulement discuter ensemble le déroulement prévu pour chaque situation-problème mais à le préciser sur papier. En effet, lors des premiers entretiens téléphoniques nous avons décidé ensemble d'analyser chaque déroulement à partir d'un premier document où nous exprimerions à l'enseignant nos motivations et les étapes prévues pour chaque situation-problème.

C'est ainsi que nous avons travaillé sur la base de trois documents, un pour chaque expérimentation. Ces documents, relativement proches de nos analyses *a priori* présentées dans cette synthèse, étaient donc débattus avant chaque séance.

Sur quelques difficultés

La mise en place

Une conclusion concerne les difficultés pour mettre en place ce genre de projets. Le sujet présent des singularités sur plusieurs plans :

- La dualité de la probabilité semble un sujet inconnu pour un nombre non négligeable d'enseignants.
- La notion bayésienne n'a pas une reconnaissance institutionnelle.
- Tant les types de raisonnements que les valeurs logiques manipulées semblent des sujets atypiques dans l'ensemble d'objets à institutionnaliser.
- Les décisions dans un contexte d'incertitude et les aspects interprétatifs semblent inhabituelles dans une classe de mathématiques.

Ces facteurs, parmi d'autres, rendent complexe l'insertion de la dualité de la probabilité dans une classe ordinaire des mathématiques. De plus, les enseignants semblent trouver dans ce genre d'expérimentations une augmentation du nombre de questions sur l'enseignement de la probabilité, plutôt que des réponses à celles ayant motivées les entretiens. Dans ce sens, la formation des enseignants autour de la dualité de la probabilité reste un sujet à approfondir. Ce dernier sujet reste inexploré dans notre synthèse, mais il s'avère fondamentale pour une mise en place satisfaisante en classe.

L'élaboration des situations-problèmes

Pendant notre recherche, notre évolution sur le plan épistémologique nous semble avoir été plus évidente que sur le didactique. Nous avons progressé dès une approche de la probabilité pouvant se caractériser comme étant classique (fréquentiste), fruit de notre formation en tant qu'enseignant, vers une autre où nous avons une vision plus intégrale (approche duale de la probabilité et de l'inférence). Néanmoins, d'un point de vue didactique, nous restons critiques envers nos expérimentions, qui nous semblent orientées par l'influence d'une approche classique sur l'enseignement de la probabilité.

Plusieurs facteurs pourraient expliquer cette "inertie" didactique, nous en retenons quatre :

- Nos pratiques en tant qu'enseignant de mathématiques au lycée pendant plus de dix ans. Ces pratiques, non significativement différentes à celles des enseignants français, se sont projetées sur nos réflexions en nous conduisant à proposer des exercices proches de ceux typiques de lycée.
- Notre place en tant qu'étudiant étranger. Cette situation relativement particulière s'est mise en évidence lorsque nous traversions l'étape de dessin de nos situations-problèmes, surtout après plusieurs mois sans trouver un enseignant pour nos expérimentations. Notre souci à que notre formation et pratiques nous mènent inconsciemment à dessiner des situations pour des scénarios plausibles chez nous (Argentine) mais non pas en France, nous renvoyait à réfléchir sur des situations proches aux typiques du système français : jeux de hasard, concentration sur des aspects calculatoires, des contextes historiques, etc. A cet effet, les pré-expérimentations ont été de grande utilité pour écarter toute différence culturelle significative, néanmoins, le souci pour ne pas nous éloigner des habitudes françaises se présentait assez fréquemment.
- Le temps de mûrissement des concepts. La séquence de cette synthèse (d'abord une enquête épistémologique, puis une analyse de manuels et finalement les expérimentations) ne s'est pas produite chronologiquement sans intersection. Les conclusions et les concepts tirés d'un chapitre ne précédaient pas forcement le début des activités du suivant. Il nous est arrivé qu'une idée se stabilisait après l'avoir investie dans notre recherche, de cette manière, la conceptualisation suivait son application. Cette conceptualisation tardive nous a mené quelques fois à appliquer des concepts avant qu'ils soient stabilisés. Par exemple, un concept ayant mûri après sa mise en place dans les expérimentations est celui des ensembles de référence. Dans ce sens, nous regrettons ne pas l'avoir bien utilisé en proposant, à la place du problème des circuits (deuxième expérimentation), une situation à des ensembles de référence multiples, tel que l'exemple de Camille, où elle évaluée sa probabilité de réussite à son examen par différentes ensembles de référence.

- La nécessité d'effectuer nos expérimentations. Notre approche par rapport aux expérimentations a évolué au fur et à mesure le temps s'écoulait sans trouver une classe où effectuer nos expérimentations. C'est ainsi que quelques idées ont été écartées pour trouver d'autres d'acceptation plus évidente. Par exemple, nous avons écarté une situation autour d'un test sur la proportion de CD abîmés dans un lot de cent CD acheté sur Internet.

Enfin, ces facteurs, et d'autres, semblent nous avoir influencé à choisir des situations dont aujourd'hui, à la fin de notre synthèse, nous ne sommes pas entièrement satisfaits, même si l'objectif majeur d'explorer les possibilités de sensibilisation de la dualité nous paraissent assez bien accompli. En fait, nous placerions nos situations-problèmes dans une sorte d'équilibre entre l'état de développement de nos idées au moment de leur élaboration et le champ de possibilités de traitement de ce sujet si particulier ; aujourd'hui, le premier nous semble avoir évolué significativement.

Les variables didactiques

Les valeurs

De reproduire les expérimentations, nous changerions les valeurs de certaines variables didactiques. En particulière, pour la troisième expérimentation, le nombre de retournements à effectuer pendant que les élèves représentent les probabilités par jetons devrait augmenter. En effet, l'information tirée des deux retournements effectués ne suffit pas à créer la tension nécessaire pour questionner le paradigme déterministe des élèves ayant choisi le principe de raison insuffisante. Il nous semble qu'en augmentant les retournements, l'information accumulée forcerait les élèves à reconsidérer l'approche entreprise, en cherchant une autre où cette information soit prise en compte.

La fonction

Les valeurs choisis pour les variables didactiques visent acheminer la résolution des problèmes par une voie déterminée. Pour certaines variables, les effets de chaque valeur possible sont plus évidents que pour d'autres. Par exemple, pour la troisième expérimentation, les effets de montrer ou non le contenu de l'urne semblent assez évidents sur les possibles procédés des élèves. Pour le premier cas, les élèves évalueraient leurs probabilités par référenciation, pour le deuxième cas, l'effet cherché est l'application du principe de raison insuffisante. Néanmoins, pour d'autres variables, leur fonction dans le problème devient moins évidente à déceler, en particulier le choix d'un contexte réel ou non.

Cette variable didactique a été testée lors de nos pré expérimentations, c'est là que nous nous sommes inclinés par les contextes réels. A l'occasion, nous avions fait l'hypothèse qu'un contexte réel ne serait pas équivalent à un autre simulé. Autrement dit, dans une étape de formation du concept de probabilité, un contexte réel placerait les élèves en situation d'incertitude plus efficacement qu'un contexte simulé ou même évoqué.

Cette hypothèse a été confirmée de manière fortuite dans une de nos pre expérimentations. Dans une classe de Terminale scientifique, nous avions mis en place une situation-problème du type jeu de hasard. En résumé, à l'occasion, les élèves, en jouant avec de pièces de monnaie, expérimentaient sur les caractéristiques du jeu au même temps qu'ils testaient des possibles solutions. A la fin de la séance, rares étaient les groupes ayant réussi à résoudre le problème. Nous avons proposé à l'enseignant de finir le problème la séance suivante. L'enseignant, en acceptant volontiers de nous céder une séance additionnelle, revint sur une question que nous avions contourné lors de la première séance : il nous a demandé s'il était possible qu'une telle suite soit effectuée sur ordinateur, vu que le programme assigne un rôle important aux simulations. Trouvant là une occasion pour tester notre hypothèse de non équivalence, nous lui avons répondu affirmativement.

Lors de la séance suivante, les élèves se trouvaient placés devant des ordinateurs où une simulation sur une feuille de calcul "effectuait" les tirages à leur place. La réaction des élèves fut immédiate et unanime : pour eux, la simulation n'était pas « la même chose » que les pièces de monnaie, nous pourrions résumer la position des élèves comme étant de non acceptation de la substitution des lancers des pièces par la simulation. Après quelques échanges, les élèves ont accepté de continuer le travail sur la feuille de calcul mais sans cacher leur mécontentement, cette acceptation pouvant s'expliquer par la pression exercée par nous et l'enseignant voie le contrat didactique.

Bien que la réaction des élèves devrait être analysée plus en détail, nous la considérons comme une confirmation de notre hypothèse de non équivalence entre contexte réel et simulé. Par son importance, les études sur les conditions d'une éventuelle équivalence méritent d'être approfondies, de la même manière que celles sur les implications de la non acceptation, ces études feront partie de nos travaux *a posteriori* cette thèse.

Le temps

La non équivalence entre contexte réel et simulé n'est pas la seule raison pour choisir le premier. Une autre concerne la qualité de l'information. Avec cette dernière expression nous souhaitons représenter une sorte de confiance que l'on porterait sur la véracité des données. En d'autres termes, une donnée douteuse serait de mauvaise qualité, de la même manière qu'une donnée constatée serait de qualité maximale.

La qualité des données intéresse la probabilité bayésienne (Gärdenfors et al., 1988). En effet, le doute sur une donnée peut, à juste titre, se transformer en un changement de une probabilité. De cette manière la valeur de la probabilité se correspond avec son signifié, un degré de certitude. Par exemple pour la deuxième séance, si le fonctionnement de chacun des composants avait été testé à l'avance par l'enseignant et juste présenté lors de la séance, sous la forme d'une liste ; les élèves pourraient avoir douté de la véracité de cette information. Ainsi, ces doutes auraient mené à penser que l'enseignant cachait des intentions didactiques avec les proportions déclarées. Si, sous ces conditions, l'on demande aux élèves de déterminer la probabilité de fonctionnement de chaque montage, il serait d'attendre que quelques-uns intègrent leurs doutes à leur évaluation de la probabilité. De cette manière, ils pondéraient la proportion de cas favorables et possibles par une sorte de coefficient de confiance. Par exemple, si la proportion de fonctionnement d'un montage est évaluée, par référenciation, en 2/3, et si les élèves considèrent l'information peut fiable, ils pourraient baisser (ou monter) la proportion 2/ 3 à une valeur accord à leurs attentes.

Cette approche, loin d'être une erreur, tient compte du caractère subjectif de la probabilité. D'ailleurs, cette subjectivité a été repérée chez les élèves lors de l'évaluation d'une probabilité bayésienne dans la troisième expérimentation, lorsqu'ils évaluaient les premières distributions *a priori* (entropie, etc.). Pour l'expérimentation du problème des circuits nous avons choisi de bloquer tout autre source d'évaluation autre que la proportion dans l'ensemble, en garantissant ainsi une qualité maximale des données.

Une remarque. Les recherches sur les conceptions des élèves autour de la probabilité ne nous semblent pas toujours tenir compte de la nature de l'objet, des raisonnements et des critères d'évaluation mises en place. Quelques-unes même, nous semblent analyser les réponses des élèves d'un point de vue fréquentiste tandis que la question posée est de nature bayésienne et pour autant de caractère subjective. Dans certains cas, ce serait plutôt les conceptions des chercheurs qui se reflètent sur les supposés diagnostiques des conceptions des élèves.

Les éléments caractéristiques proposés dans cette synthèse nous semblent mettre en évidence la complexité des mécanismes associés à l'évaluation d'une probabilité en fonction du type d'objet. Ces éléments contribueront à la recherche sur les représentations des élèves.

Les raisons précédemment énoncées en faveur de l'utilisation de contextes réels (qualité de l'information et non équivalence) ne sont pas les seules. Une autre, et encore une fois pour la probabilité bayésienne, concerne la fixation du type d'objet. En d'autres termes, la présence matérielle des objets nous semble contribuer à cerner la nature de l'objet et le caractère épistémique d'une probabilité bayésienne. Par exemple, avec la présence des montages lors de l'expérimentation des circuits nous avons cherché à réduire ou même enlever les ambiguïtés possibles concernant la nature de l'objet. En effet, il y avait d'une part l'enseignant qui parlait d'un objet en particulier (un montage), et d'autre part, cet objet présent devant la classe qui renforçait la nature de la question.

Les réactions des élèves semblent confirmer que la variable "présence de l'objet" a contribué à fixer sa nature, personne n'a proposé la reproduction de l'épreuve, ni s'est référé à ce qui arriverait à long terme si on la reproduisait.

Le contexte réel est un choix coûteux en temps investi avant et durant la séance. En effet, la préparation des objets demande un effort considérable, et même, en étant soigneux dans leur mise au point, on ne peut pas garantir leur fonctionnement lors de la séance. L'investissement en temps ne se conclut pas lors de la mise au point des objets ; au temps de préparation, nous devons ajouter le temps didactique, celui investi cette fois-ci durant la séance.

Les trois expérimentations ont demandé ces deux temps d'investissement. Néanmoins, de ce point de vue, l'expérimentation la plus critique a été la deuxième, celle des circuits. La préparation des composants, la rupture intentionnelle de pièces pour arriver aux proportions attendues, le branchement des composants, etc., enfin, tout ce travail de préparation nous semble impraticable pour l'enseignant d'une classe ordinaire. A ce temps, il faut ajouter celui du test, lors de la séance, où chaque groupe détermine le fonctionnement des composants, ce temps, est aussi précieux.

Un risque pour ce genre de situations concerne la substitution de la production des données par son énonciation ou sa simulation. Cette substitution permettrait à l'enseignant d'épargner le temps de production des données et celui de préparation, mais le coût retomberait sur la qualité de l'information, la fixation de l'objet, et la non transposition des conclusions.

Dans ce sens nous devons signaler le total respect de l'enseignant envers nos propositions, quelques-unes non accompagnées de leurs justifications. Néanmoins, et même si nous n'avons perçu aucun signe de malaise de sa part pour nos choix didactiques coûteux en termes de temps, nous admettons que la réalisation dans une classe ordinaire devient très improbable. De fait, aucune de nos situations-problèmes n'a été conçue en visant sa reproduction dans une classe. Nous avons juste tenté d'explorer quelques conditions pour une sensibilisation à la dualité de la probabilité. En particulièr, le problème des circuits nous semble sans intérêt majeur pour un enseignant. Malgré son apparente pauvreté, nous avons tiré des conclusions significatives concernant les possibilités d'une sensibilisation à la notion bayésienne de type épreuve générique, en particulier ce qui concerne la fixation de la nature de l'objet, les critères d'évaluation et la qualité de l'information.

En fait, nos expérimentations doivent se placer dans un contexte assez particulier de la didactique des mathématiques, où la notion bayésienne était, inconnue pour les uns, juste une subtilité banale pour les autres, et même certains étaient sceptiques devant la possibilité d'un traitement dual de la probabilité en classe.

Toutes ces impressions réactives au sujet ont nourri nos réflexions. Par exemple, la deuxième expérimentation (les circuits) prouve qu'il est possible a) le traitement de la notion bayésienne de la probabilité en classe et b) de contrôler l'émergence des critères d'évaluation pour une épreuve générique. Nous caractériserions notre recherche donc, comme une sorte d'exploration des voies possibles pour le traitement de la dualité de la probabilité.

Nous devons signaler que les réservés constatées devant la question de la possibilité de traitement de la dualité de la probabilité en classe ne sont pas propres au contexte français. Quelques chercheurs américains sont déjà exprimés dans la même direction (Moore, 1997). Néanmoins, ces dernières années, la communauté didactique internationale semble plus optimiste, en Espagne de même qu'un Italie, des équipes travaillent en intégrant dans leurs sujets de recherche les deux approches de la probabilité (Batanero, Diaz, & de la Fuente, 2005; Garuti, Orlandoni, & Ricci, 2008; Schield, 2001).

Dans ce sens, il nous semble que les expérimentations réalisées en BTS montrent qu'il est possible une sensibilisation à la dualité de la probabilité et qu'il nécessaire réfléchir à sa reconnaissance institutionnelle. Reconnaissance qui devrait conduire dans un deuxième temps à s'intéresser à la formation des enseignants. Mais que, dans aucun cas, la dualité de la probabilité est un sujet impossible ou trop complexe pour les élèves.

Néanmoins, nous restons prudents devant la possibilité du traitement de la dualité de la probabilité dans une classe ordinaire, au moins avec l'orientation actuelle des programmes. En effet, le programme de la classe de Première ont entrepris une voie bien définie par rapport à la probabilité. Premièrement, il ne retient qu'une seule approche de la dualité, deuxièmement les conseils s'orientent vers les simulations, au lieu de la manipulation réelle, et troisièmement, il n'inclus pas des notions de net intérêt statistique (vraisemblance, critères décisionnels, principes d'évaluation, etc.).

Pour ce qui concerne ce dernier point, l'introduction de la dimension interprétative de la probabilité à l'enseignement entraîne la reconnaissance institutionnelle d'un nombre non négligeable de concepts statistiques. La liste ne se limite pas aux deux interprétations de la probabilité, et si de plus, nous intégrons une approche décisionnelle, alors le nombre de concepts à introduire en classe requiert une analyse didactique minutieuse qui dépasse les limites de notre recherche. Étudier quels objets doivent acquérir le statut institutionnel et quels peuvent rester sous-entendus, nous semble fondamentale pour l'enseignement de la statistique, en particulièr l'inférentielle.

L'approche des programmes, étant fortement prise en compte pour les enseignants de mathématiques, fait de la dualité de la probabilité, un sujet difficile à rendre sensible en classe. Les enseignants semblent soumis à une forte pression concernant l'application non seulement des concepts précisés dans le programme mais aussi les recommandations qui les accompagnent. De plus ils doivent se confronter à leurs propres représentations sur l'enseignement des mathématiques.

Par exemple, pour ce qui concerne les contextes réels, d'une part, les entretiens avec les enseignants de nos pré expérimentations et ceux effectués au BTS, nous renseignent d'un fort intérêt pour les contextes simulés. Il semble qu'un objet matériel est mieux perçu pour des niveaux plus élémentaires de l'enseignement; au lycée, les enseignants associeraient les mathématiques à des entités abstraites ou symboliques. D'autre part, les indications des programmes, propositions de manuels et quelques recherches locales en didactique, font de la manipulation d'objets, un sujet rare ; la tendance étant largement vers la substitution de la réalité par des simulations. De cette manière, lorsque l'enseignant adhère à une telle substitution, il obtiendrait non seulement l'approbation de ses collègues mais aussi celle de la communauté didactique en générale.

Dans toutes nos expérimentations et même celles les ayant précédé, nous avons été confrontés à l'intérêt des enseignants pour introduire les ordinateurs. Les contextes réels ont une forte concurrence premièrement par le choix professionnel d'associer les mathématiques à d'entités abstraites et deuxièmement par la pression du système qui les promeuve constamment.

Nous exprimons nos réserves, cette fois-ci devant l'utilisation des ordinateurs pour certains concepts, ou au moins lors de la construction de la notion de probabilité. En effet, il nous reste le doute que les conclusions tirées lors d'un travail dans des contextes évoqués et simulés soient, pour les élèves, transposables à la réalité qui les entoure.

Les registres de représentation

Dans cette synthèse nous avons repéré deux dimensions pour la probabilité, l'une sémantique l'autre opératoire. La première réunit les aspects interprétatifs de la probabilité, la deuxième les calculatoires. Nous avons aussi proposé un niveau hiérarchique pour ces deux dimensions, le premier serait de degré supérieur au deuxième, en d'autres termes, l'interprétation de la probabilité déterminerait le genre de transformations plausibles à effectuer sur une valeur d'une probabilité.

A chacune des dimensions nous avons associé aussi un registre sémiotique, les aspects interprétatifs se développeraient sur le registre langagier, les calculatoires sur le numérique et symbolique, ce dernier donnant les règles d'usage et de transformation à toute expression numérique de la probabilité.

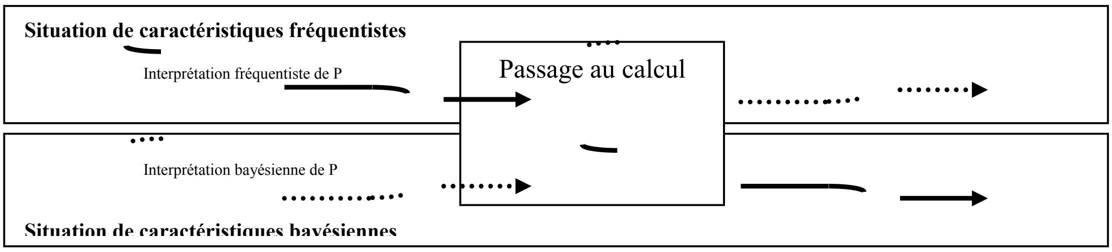
Une conclusion (ou plutôt une remarque) concerne la relation entre les dimensions sémantique et opératoire, ces dimensions se complémentent sans se substituer. Le champ du possible pour l'une, n'est pas empruntable à l'autre. Ce que nous pouvons exprimer sur le registre de la langue n'est pas possible sur le numérique et vice-versa, le degré de précision et les critères de transformation fournis sur les registres numérique et symbolique ne sont pas possibles ailleurs.

De cette relation de complémentarité sans substitution entre les dimensions sémantique et opératoire, nous tirons une conclusion d'ordre didactique. Si lors de la résolution d'un problème, un élève réussit à effectuer correctement ses calculs, cela n'implique pas qu'il compris son signifié. En d'autres termes, nous ne trouverons pas des traces du signifié attribué à une probabilité juste en regardant la résolution numérique d'un exercice. Pour cela, nous devons tourner notre attention vers le registre langagier, où une interprétation trouve les outils nécessaires à un tel genre d'expression.

Le schéma de représentation théâtral évoqué au début du Chapitre III vise à mettre en évidence d'une manière figurée l'espace des possibilités pour chaque registre sémiotique ainsi que leurs relations. De cette manière, si nous nous intéressons aux aspects interprétatifs de la probabilité, nous serions obligés de considérer un support linguistique (orale ou écrit), toute intention de précision en termes de valeurs, requerra des expressions dans les registres numériques et symboliques.

Chacun de ces deux registres est plus approprié que l'autre en fonction de la dimension considérée. C'est ainsi que toute transformation sur le numérique, registre unifié pour les deux interprétations de la probabilités, ne retiendra pas le signifié attribué à la valeur en question. Pour savoir son signifié postérieurement à la transformation, en fait pour le récupérer, il faudra chercher le signifié assigné avant la transformation numérique. Le signifié serait donc conservé a travers les transformations numériques. De cette manière, une probabilité dans un contexte donné peut changer de valeur mais pas de signifié. Une erreur courante réside à attribuer un signifié à un calcul et puis lors d'une transformation à lui en attribuer un autre, cette erreur a été représentée dans la figure que nous reproduisons ci-contre.

Erreur par échange d'interprétation postérieur aux calculs



Dans ce sens, le contexte joue un rôle important lorsque, après une transformation numérique, on souhaite retrouver le signifié d'une probabilité. Le contexte est une source de (re)construction de sens, de repérage d'indices permettant la récupération de l'interprétation d'une probabilité. A cet effet, plusieurs auteurs ont attiré l'attention sur des erreurs à cause de l'échange d'interprétation commise après une transformation numérique. Par exemple, une probabilité précédant une transformation est interprétée en termes fréquentistes et celle d'après comme étant bayésienne. Plusieurs travaux pointent sur cette erreur récurrente, principalement lors de l'interprétation d'une test d'hypothèse fréquentiste (Jaynes, 1976; Lecoutre, 2005; Régnier & Oriol, 2001; Schield, 1997).

Une autre observation concerne les difficultés, associées à la dualité, par l'effort additionnel que représente le travail à déployer dans le registre de la langue, travail qui donne lieu au développement de l'interprétation d'une probabilité. En d'autres termes, non seulement on doit rester attentif aux aspects calculatoires (registres numériques et symboliques) mais aussi interprétatifs (registre langagier). Ces derniers, avec les difficultés propres à ce registre, telles que polysémies, ambiguïtés, etc. Un risque pour un travail en classe sur les interprétations consiste en la standardisation de la terminologie par la voie de l'utilisation de mots clefs, de cette manière élèves et enseignants appliqueraient des automatismes basés sur des indices du problème pour trouver le mot clef associé.

Dans nos expérimentations, cette standardisation n'a pas été détectée, d'une part parce que l'enseignant était au courant de notre projet, devenu en grande mesure le sien, d'autre part parce que le nombre d'expérimentations a été réduit. Nous devrions nous attendre à ce qu'une telle standardisation par des mots clefs se produise lorsque le nombre d'exercices s'accroît considérablement.

Les objets à institutionnaliser

Par contre, nous avons observé dans nos expérimentations que les conclusions et les institutionnalisations concernant les aspects interprétatifs et décisionnels, n'ont pas dépassé le plan du registre oral, aucune trace n'est restée de ces concepts sur les cahiers des élèves. Cela pourrait s'expliquer par leur non appartenance à l'ensemble de concepts ou contenus du programme.

Néanmoins, l'institutionnalisation de la dualité en classe nous semble poser des problèmes. La question de la manière, dans l'état actuel de l'enseignement, d'institutionnaliser et d'évaluer nous semble difficile à résoudre. Ces objets, par les caractéristiques du registre dans lequel ils s'expriment pourraient risquer de se réduire à des mots clefs. En effet, pour surmonter les ambiguïtés propres au registre, professeurs et élèves pourraient faire recours à des termes standardisés, cherchant ainsi à faciliter le repérage de ces objets lors d'une évaluation. Le risque de la standardisation ne concerne seulement pas seulement les interprétations, en principe tous les objets représentés dans le registre langagier seraient confrontés à ce genre de difficultés y compris les principes d'évaluation et les critères décisionnels.

Dans l'état actuel de notre recherche, il nous est impossible de dissocier interprétation de la probabilité et sa fonction comme outil de prise de décision. Cela ne veut pas dire que nous ne concevons que des situations décisionnelles pour l'enseignement de la probabilité, mais elles doivent faire partie de l'ensemble de situations proposés aux élèves.

La question de l'institutionnalisation d'objets propres à la discipline ne nous semble pas anodine, en particulier par le caractère consensuel de quelques-uns (principes d'évaluation et critères décisionnels). Par exemple, un des critères décisionnels les plus élémentaires reste celui de la maximisation de l'espérance. Ce critère requiert, par définition même, un consensus au sein de la classe pour être accepté. En d'autres termes, un critère, n'est pas déductible, il ne peut être qu'accepté et cela par les poids et la cohérence des arguments, de même que les principes d'évaluation d'une probabilité bayésienne.

Dans ce sens, les nouveaux objets de statistique, les statuts logiques (caractère consensuel) de quelques-uns et la prise en compte de la dimension sémantique nous interrogent dans deux directions, d'une part vers l'étude des conditions de l'enseignement de chacun d'eux et d'autre part vers le champ des théories didactiques modélisant cette problématique. Pour ce qui concerne la deuxième direction, nous confirmons, à la fin de cette synthèse notre position de prudence devant des théories importées de la didactique des mathématiques. En effet, il continue à nous paraître prématuré de retenir une théorie déterminée. Nous ne sommes pas en conditionnes d'assurer qu'une théorie didactique provenant des mathématiques puisse modéliser le processus d'enseignement et d'apprentissage de la statistique.

Pour cela, notre position peut être caractérisée comme prudente devant l'utilisation d'un cadre théorique donné. En effet, le sujet dualité de la probabilité nous semble d'une relative nouveauté dans le champ de la didactique des mathématiques, cela étant, il nous semble risqué d'adhérer à telle ou telle théorie didactique lorsque nous ne nous sommes pas suffisamment approprié des enjeux des objets épistémologiques en question ; l'analyse de l'adaptabilité des théories didactiques des mathématiques à celle de la statistique peut être considère un autre sujet des perspectives suivant notre recherche.

Connaissance et méconnaissance

Pour nos situations-problèmes nous avons considéré l'aspect décisionnel de la probabilité, c'est ainsi que nous l'avons placé comme un outil rationnel d'aide à la décision. Cette approche a fait que nos expérimentations ont demandé aux élèves des argumentations. Ce genre de demande les a placés de manière alternée en positions de connaissance et de méconnaissance. La première lorsqu'ils étaient en train de trouver une solution aux problèmes, les deuxièmes lorsqu'ils les devaient les justifier.

Dans nos expérimentations s'est produite une sorte de partition de ce genre d'activités, les élèves ont abordé les premières en relative autonomie et les deuxièmes uniquement sous la coordination de l'enseignant. Cette séparation pourrait s'expliquer par l'effet d'au moins trois facteurs :

- Le premier concerne les difficultés propres aux tâches métacognitives, la réflexion et la description à autrui, de ses propres actes, est reconnue comme étant d'une difficulté supérieure à l'activité cognitive (Flavelle, Miller, & Miller, 1993). Les élèves, lorsqu'ils devaient justifier leurs choix, auraient été confrontés à des tâches significativement plus complexes que lorsqu'ils ont effectué des calculs, de cette manière, la tutelle de l'enseignant aurait un effet compensateur de cette difficulté.
- Le deuxième facteur concerne une répartition de tâches qui s'installerait au sein de la classe, répartition devenant élément du contrat didactique. En effet, il pourrait exister, et pour des raisons non seulement cognitives (maintien de la discipline en classe, gestion du temps, etc.) une répartition de tâches au sein de la classe, pour laquelle, les élèves en effectueraient certaines de manière autonome et d'autres, sous la guidance de l'enseignant, dans ce cas, les cognitives à la charge des élèves, les métacognitives sous la tutelle de l'enseignant.
- La troisième explication concerne le caractère consensuel de quelques concepts. Nous avons fait référence à une caractéristique de certains concepts de statistique, caractéristique qui serait une conséquence du paradigme indéterministe. Nous avons mentionné que la validation de certains concepts ne s'effectue que par consensus, où leur poids ne repose pas sur la valeur logique des conclusions tirées mais sur la cohérence des argumentions. Les exemples sont nombreux et cette caractéristique est vérifiable même sur des notions bien élémentaires. Des critères tels que la maximisation de la probabilité, celui de maximisation de l'espérance et le principe de raison insuffisante sont des exemples où un consensus est indispensable à leur application.

Modifications des situations-problèmes

Lors de la conception des situations-problèmes nous avions considéré la possibilité d'une suite pour chacun des trois problèmes. Ces suites, d'une part, retenaient les éléments qui pour des raisons diverses ne pouvaient pas être traités lors de la séance, et d'autre part, elles constituent une réserve de paramètres nous permettant d'adapter la situation originellement proposée à l'enseignant si besoin était. Pour chacune de nos trois expérimentations il y a un ensemble de modifications possibles, où le contrôle s'effectuait par des variables didactiques non exploitées lors des expérimentations.

Première situation-problème

A nos yeux, la modification la moins intéressante des trois. Son expérimentation demanderait probablement une séance entière. La modification à la première situation-problème consistait à, éventuellement dans une deuxième séance, poser la même question que lors de la première (avec combien de jetons convient-il de jouer ?) mais en substituant les jetons par des dés. En principe, une telle modification n'entraînerait que des changements au niveau de la dimension calculatoire de la probabilité.

En effet, la notion de la probabilité sous-jacente à cette nouvelle situation reste invariante, elle est de nature fréquentiste, la probabilité porte sur la série infinie des hypothétiques lancers effectués, cette fois-ci avec des dés. Même le critère décisionnel ne diffère pas de celui de la première situation-problème, il reste le critère de maximisation de l'espérance. L'application de ce critère donne comme réponse optimale deux choix possibles, soit on joue avec cinq dés, soit on joue avec six dés, pour tout autre choix, l'espérance de gain est inférieure.

La dimension interprétative y compris les aspects décisionnels étant les mêmes, la seule différence consisterait en une augmentation non négligeable des difficultés de l'ordre calculatoire.

Nous avons présenté cette modification, ou plutôt cette suite, à l'enseignant, s'il souhaitait renforcer les notions abordées lors de la première situation-problème. Cette modification, n'a pas été expérimentée finalement. Lors d'un de nos entretiens ayant suivi la première expérimentation, l'enseignant nous avait confié avoir donné cette suite comme devoir à la maison. Nous ne disposons d'aucune trace de l'éventuel travail des élèves.

Deuxième situation-problème

Les modifications à cette deuxième situation sont plus en accord avec le sujet que nous intéresse, que celles de la première situation. Cette fois-ci, elles concernaient les valeurs de probabilité des composants des montages. Cette suite n'implique pas forcement une nouvelle séance, elles avaient était préparées si l'on disposait de plus de temps à la fin de l'expérimentation.

Cette suite consistait à revoir les valeurs des probabilités des montages par la vérification du fonctionnement d'un ou plusieurs de leurs composants. Plus précisément, une fois ayant résolu le problème en proposant un montage comme étant celui de plus haute probabilité de fonctionnement, l'enseignant prendrait un montage et il testerait le fonctionnement d'un de ses composants pour finalement poser la question : avec ce nouveau test que je viens d'effectuer, choisirez-vous le même montage ?

Les modifications des valeurs des probabilités bayésiennes du fonctionnement des composantes passeraient de l'intervalle ouvert (0,1) à une des valeurs de certitude 0 ou 1, ces modifications auraient des conséquences sur les valeurs des probabilités, aussi bayésiennes, du fonctionnement des montages. Des modifications qui pourraient, à leur tour, reconsidérer la décision prise lors de la première partie du problème.

Il nous semblait intéressant d'observer les éventuels débats qu'une telle suite pourrait provoquer chez les élèves. Principalement parce que l'enjeu est nettement interprétatif avec un travail restreint sur le plan calculatoire. Nous cherchions avec cette suite à provoquer des échanges autour du signifié de la probabilité bayésienne et de sa valeur en tant que mesure de certitude. De plus, une telles suite, aurait permis à l'enseignant, lors de l'institutionnalisation, de renforcer le caractère épistémique de la probabilité bayésienne.

Cette suite, et vu que nos prévisions nous suggéraient la non disponibilité de temps supplémentaire, fut juste commentée à l'enseignant lorsque nous discutions l'analyse *a priori* de la deuxième situation-problème. Cette prévision s'est confirmé lors de l'expérimentation, rappelons que la séance du problème des circuits fut conclue sans la phase d'institutionnalisation pour des raisons de temps.

Troisième situation-problème

Pour la troisième expérimentation nous avions souhaité profiter des avantages de la vitesse de calculs de l'ordinateur pour concentrer l'attention de la classe sur la dimension sémantique de la probabilité et sa fonction comme outil décisionnel. Quelques variations au problème original avaient été déjà introduites dans le questionnaire, parmi d'autres, analyser les effets sur les probabilités lorsque l'on permute les couleurs des billes observées et prendre la décision avec un nombre plus réduit de retournements.

En particulier pour cette situation-problème nous avions envisagé une suite autour de deux axes, l'un sur les critères d'évaluation, l'autre sur le signifié de la probabilité en tant que degré de certitude, toutes les deux articulées à partir d'une seule modification.

La modification envisagée consistait à dévoiler le contenu de l'urne après avoir rempli le questionnaire, et puis analyser de quelle manière cette nouvelle information pourrait être intégrée à la démarche effectuée. Une réaction possible des élèves, souhaitée en fait, consiste à tourner l'attention vers la première distribution *a priori*, celle évaluée par le principe de raison insuffisante. Ce principe avait était utilisé à cause du manque d'information concernant l'urne, vue que l'on ne connaissait rien sur elle, on n'avait pas des raisons pour privilégier une composition sur une autre, la distribution était donc équirépartie.

Mais maintenant, avec ce dévoilement, les élèves auraient une raison pour croire plutôt dans les unes que dans les autres, cette raison serait proportion utilisée dans l'ensemble. En termes de nos éléments caractéristiques, ils évalueraient les probabilités de chaque composition par référenciation.

Un tel procédé n'arriverait pas à être une référenciation *stricto sensu*, pour cela, les élèves devraient pouvoir déterminer avec précision toutes les combinaisons possibles pour chacune des compositions et en déduire leurs proportions dans l'ensemble. Ces calculs restent évidement inabordables pour ces élèves avec leurs connaissances. En tout cas, ce qu'ils pourraient faire serait une estimation de cette proportion. De cette manière, et par une approximation numérique, ils évalueraient leur attente pour chaque possible composition.

Cette nouvelle première distribution reflèterait donc leurs degrés de certitude cette fois-ci basés sur la proportion des billes contenues dans l'urne. Les distributions possibles pour la première série *a priori* seraient nombreuses, principalement à cause de la nature imprécise de sa détermination. Les valeurs varieraient donc d'équipe en équipe, mais le principe devrait rester invariant, plus proche de la proportion de l'urne est une composition, plus probable elle devient.

Rappelons que nous avons choisi une composition de soixante dix pour cent de billes orange et trente pour cent de noire. De cette manière, les compositions les plus probables deviendraient la C (NNOO) et la D (NOOO), vu que la proportion pour cette dernière est plus proche de celle de l'urne, on devrait s'attendre à que les élèves lui assignent une plus forte probabilité. Les autres compositions devraient avoir assignées des probabilités plus faibles, en fonction de leurs distances à la proportion de l'urne.

La représente les successives distributions de probabilités pour la même série de retournements mais cette fois-ci, avec une première distribution *a priori* révisée par l'information fournie par l'enseignant. Lors des dix retournements, et en appliquant la formule de Bayes de manière itérée, on peut observer que les compositions C et D restent si proches que l'on pourrait considérer toutes les deux comme étant de même probabilité.

Evolution des distributions de probabilités (révisée)

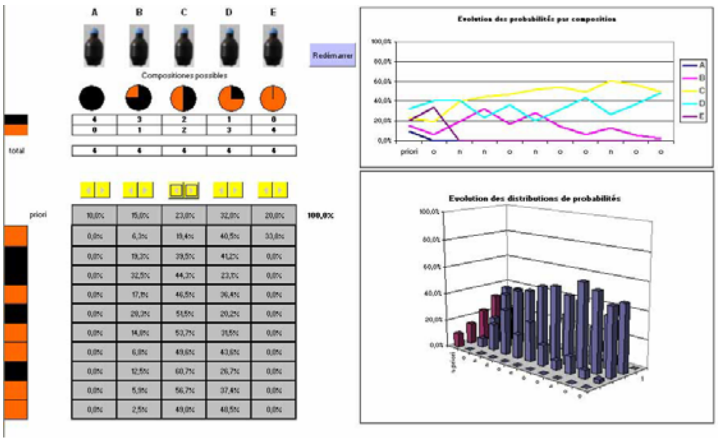


Figure 49

Cette modification de la situation-problème originale nous semblait potentiellement très riche en débats. Elle met en évidence le caractère subjectif de la probabilité et sa dépendance à l'information disponible, en même temps qu'elle permet de réfléchir sur les conditions d'application des différents critères d'évaluation. En fait, la feuille de calcul fut conçue pour permettre de jouer librement sur les deux ensembles de paramètres du modèle, d'une part les premières distributions *a priori*, d'autre part les couleurs des billes. Le premier ensemble de paramètres est rapidement modifiable par les boutons jaunes, le deuxième, par les lettres "o" ou "n".

Cette suite nous semblait riche pour les débats qu'elle pouvait provoquer autour de l'interprétation de la probabilité bayésienne. De sa réalisation, nous attendions tirer des données importantes concernant non seulement les représentations et raisonnements des élèves, mais aussi sur les possibilités d'institutionnalisation du caractère subjectif de la probabilité. Cette suite, de même que celles des autres expérimentations, n'a pas eu lieu, une explication immédiate peut se trouver dans le manque de temps lors de l'expérimentation de la troisième situation-problème, rappelons que la séance se termine avant même d'avoir fait le bilan des réponses au questionnaire.

D'ailleurs, les conditions dans lesquelles se sont déroulées nos expérimentations nous semblent un facteur ayant limité les possibilités de tester les suites envisagées. Les contraintes qui pèsent sur le système, la non reconnaissance institutionnelle du sujet, notre état de conceptualisation au moment de nos expérimentations et nos possibilités de communication dans la langue française à ce moment de notre recherche, tous ces facteurs nous semblent avoir influencé négativement et limité les possibilités de suites à nos expérimentations. Les deux derniers facteurs sont non négligeables, même si nous ne disposons pas d'indices pour le confirmer. C'est en tout cas par les non dits, les silences et notre impression de manque d'outils linguistiques précis lors de nos entretiens, que nous nous inclinons a considérer ces deux éléments comme des facteurs contribuant à limiter les possibilités d'avoir effectuer ces suites.

Par exemple, lors d'un des derniers entretiens avec l'enseignant du BTS, nous avions renoncé à une explication jugée trop complexe (dans l'état d'avancement de notre conceptualisation et maîtrise linguistique du moment). Cette explication concernait les différences entre les deux interprétations de la probabilité, des différences que nous souhaitions pointer sur la troisième expérimentation. Pour les mettre en évidence, quelques modifications sont nécessaires à cette troisième situation-problème, que nous essaierons de détailler ci-contre.

Le troisième problème, si l'on se concentre sur la probabilité portée sur la bouteille, répond aux conditions du paradigme bayésien. En effet, la question est de nature épistémique et les critères d'évaluation de la première distribution de probabilités *a priori* sont selon les cas, le principe de raison insuffisante (situation originale) ou une estimation de l'ensemble fini de référence (situation modifiée). Pour passer au paradigme fréquentiste, quelques modifications sont nécessaires.

La première est déjà présente dans la suite que nous avons proposé, elle consiste en dévoiler le contenu de l'urne. De cette manière, en ayant des connaissances sur les hypothèses du modèle, on peut déduire la fréquence d'apparition de chaque composition possible.

La deuxième modification concerne la nature de l'objet. Elle consiste à ne plus s'intéresser au contenu d'une bouteille particulière, mais à ce qu'il arriverait si l'on reproduit l'expérience un nombre infini de fois. Ce changement de nature d'objet pourrait s'effectuer en deux étapes, la première consistant en ouvrir la bouteille devant les élèves et sous la justification de vérifier la prévision effectuée. De cette manière, la valeur logique portée sur la bouteille changerait, en passant d'une valeur dans l'intervalle (0,1) à une autre de certitude (0 ou 1). La deuxième étape, par un changement d'approche, en demandant la fréquence avec laquelle ces prévisions seraient correctes. De cette manière, l'objet ne serait plus une bouteille en particulière mais l'ensemble infini de bouteilles.

Dans ce passage de paradigme, nous voyons concourir les quatre éléments caractéristiques typiques des interprétations de la probabilité, en effet :

- nature d'objet : la situation originale (bayésienne) s'intéresse à une bouteille, la fréquentiste à une attente à long terme
- valeurs logiques : on passe d'une mesure de certitude dans la situation bayésienne à des valeurs de certitude, chaque réalisation n'admet que deux valeurs possibles, de plus, la série ne peut être que vraie ou fausse.
- types de raisonnement : la situation bayésienne s'évalue éventuellement par référenciation et abduction, la fréquentiste par déduction.
- principes de quantification : la bayésienne, selon le cas, par les principes d'équiprobabilité ou raison insuffisante, la fréquentiste par celui d'équiprobabilité.

Cet ensemble de modifications du problème, donnant lieu au changement de paradigme probabiliste, n'a pas été partagé avec l'enseignant lors de nos entretiens, la principale raison étant la non maturation de ces concepts au moment de nos entretiens.

Les perspectives sont étroitement liées aux conclusions. En effet, le caractère expérimental que nous attribuons à cette recherche nous mène à considérer nos conclusions comme n'étant que provisoires. Le sujet nous semble bien mériter une suite, vu le faible nombre de situations-problèmes expérimentées. Bien que quelques lignes de recherche ont été déjà suggérées lors de la présentation des conclusions, nous présenterons sous la forme d'une liste quelques sujets qui découlent de manière plus ou moins immédiate de cette synthèse :

- Les théories didactiques s'intéressant à l'enseignement de la statistique. Nous restons dans notre position initiale, probablement plus convaincus encore. Il nous semble prématuré s'affilier à telle ou telle théorie importée des didactique des mathématiques pour porter un regard sur l'enseignement de la statistique. Nous sommes loin encore de maîtriser les enjeux épistémologiques de base de cette discipline, la relative nouveauté d'un sujet si élémentaire comme la dualité de la probabilité semble le prouver. La poursuite donc d'éléments conceptuels visant des constructions théoriques pour aborder l'enseignement de la statistique nous semble un sujet indispensable, et cela constitue une perspective pouvant être caractérisée comme générale dans nos travaux à entreprendre dans le futur.

- La formation des enseignants. Ce sujet a resté inexploré dans notre synthèse, pour des raisons déontologiques, nous n'avons pas analysés cette facette de notre travail. En effet, la nouveauté qui représente ce sujet aux enseignants a fait qu'une partie de nos entretiens ont été consacrés à discuter ensemble sur les enjeux de la dualité de la probabilité. Il arrivait souvent que les enseignants découvraient la notion bayésienne lors de nos premiers entretiens. De cette manière, se produisait une sorte de situation d'enseignement entre le professeur et nous, où nous expliquions non seulement notre point de vue didactique mais aussi quelques aspects épistémologiques de la dualité de la probabilité. Ce double rôle auquel nous avons été confrontés nous a mené à ne pas aborder les aspects liés à la formation des enseignants, il reste donc un sujet à développer à part entière.
- Les conceptions des élèves. La prise en compte de la notion bayésienne de la probabilité ouvre l'analyse de conceptions des élèves à des nouvelles questions. Il nous semble intéressant de considérer la notion sous-jacente de la probabilité dans un problème, lors de l'étude de conceptions des élèves, cette prise en compte contribuerait à mieux observer les différentes approches personnelles mobilisés par les élèves lors de la résolution des problèmes.
- Sur le niveau de la scolarité pour une introduction à la probabilité. Etant convaincus que la statistique doit s'introduire plus tôt dans l'enseignement, nous avions souhaité effectuer des expérimentations en niveaux plus élémentaires que le BTS. Il nous reste à analyser dans quel niveau il conciendeait d'introduire la probabilité, et probablement en prenant compte non seulement l'état du développement cognitif et les connaissances nécessaires des élèves mais aussi le moment propice en fonction des contraintes du système, qui semblerait être de plus en plus réfractaire aux contextualisations dans la mesure que l'on monte de niveau de scolarité. De plus, bien que nous avons confirmé le caractère incontournable de la probabilité, on pourrait même envisager une sensibilisation progressive par niveaux. Enfin, des nombreuses questions se posent autour du niveau de l'enseignement de la probabilité.
- Les éléments caractéristiques. Cette synthèse devient une sorte de présentation pour les éléments caractéristiques visant l'identification des deux approches de la probabilité. Ils doivent être analysés par la communauté didactique, et probablement révisés et même enrichis.
- Des retombées sur des recherches autour des tests statistiques. Les interprétations de la probabilité et leurs paradigmes associés sont les piliers des méthodes inférentielles

statistiques. Les approches interprétatives de la probabilité concernent ceux qui s'intéressent à l'enseignement de notions liées aux tests statistiques. Nous attendons donc que cette synthèse leur apporte des éléments sur leurs sujets d'étude.

- Sur certaines variables didactiques. Pour nos situations-problèmes nous avons admis un ensemble d'hypothèses concernant les variables didactiques. Par exemple, nous avons admit que la présence matérielle des objets renforçait la nature des objets. Nous souhaitons approfondir des recherches dans cette direction, en cherchant à mieux connaître l'influence de ces variables dans les raisonnements des élèves. De plus, il nou semble fondamental d'examiner plus en detail les conditions d'une hypothètique transparence en simulations, de la même manière qu\analyser l'influence de la qualité de l'information lors d'une situation bayesiénne.
- Enfin, il nous faut continuer en expérimentant d'autres situations-problèmes, dans lesquelles nous pouvons réinvestir les leçons tirées de cette première mise en oeuvre de problèmes. Il ne s'agira plus alors d'étude exploratoire mais de situations plus proche des besoins des enseignants.

Bibliographie

Barbier, J.-M., & Galatanu, O. (2000). *Signification, sens, formation*.Paris: Puf.

Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances.

Bernoulli, J. (1713). *The art of conjecturing* (E. D. Sylla, Trans. 2006 ed.). Bâle: Johns Hopkings.

Brousseau, G. (1998). *Théorie des situations didactiques*.Grenoble: La Pensée Sauvage.

Brousseau, G., Brousseau, N., & Warfield, V. (2001). An experiment on the teaching of statistics and probability. *The Journal of Mathematical Behavior, 20*(3), 363-411.

Carnap, R. (1947). *Meaning and necessity. A study in semantics and modal logic*.Illinois: The University of Chicago Press.

Carranza, P. (2005). *Déplacement de problématique en l'enseignement de la statistique au lycée en france.* Paper presented at the XXXVII Journées de statistique de la Société française de statistique, Pau, France.

Condorcet, J.-A.-N. (1805). *Eléments du calcul des probabilités* (F.-J.-M. Fayolle, Trans.). Paris: Fayolle, F.-J.-M.

Crombie, A. C. (1980). *Styles of thinking and historiography of science.* Paper presented at the Sociedad Española de Historia de las Ciencias, Madrid. Espagne.

Chevallard, Y. (1985). *La transposition didactique* (1991 ed.). Grenoble: La Pensée Sauvage.

Chiasson, P. (2005). Abduction as an aspect of retroduction. *Semiotica, 153*(1-4), 223-242.

Dale, A. (1999). *A history of inverse probability: From thomas bayes to karl pearson*.New York: Springer-Verlag.

de Finetti, b. (1937). La prévision: Ses lois logiques, ses sources subjectives, *Annales de l'I.H.P.* (Vol. 7, pp. 1-68): Numdam.

De Finetti, B. (1974). *Theory of probability*:Wiley classics library.

Douady, R. (1986). Jeux de cadres et dialectique outil-objet. *Recherche en didactique des mathématiques, 7*(2), 5-31.

Droesbeke, J.-J., Fine, J., & Saporta, G. (2002). *Méthodes bayésiennes en statistique*.Paris: Sfds.

Duval, R. (1993). Registres de représentation sémiotique et fonctionnement cognitif de la pensée. *Annales de Didactique et de Sciences Cognitives, 5*, 37-65.

Feienberg, S. (2006). When did bayesien inference become "bayesian"? *Bayesian Analysis, 1*(1), 1-10.

Gärdenfors, P., Sahlin, N.-E., Ramsey, F., Luce, D., Raiffa, H., Savage, L., et al. (1988). *Decision, probability and utility*:Cambridge University Press.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). *Subjective probability*, 129-161.

Gras, R., Ag Almouloud, S., Bailleul, M., Larher, A., Polo, M., Ratsimba-Rajohn, H., et al. (1996). *L'implication statistique*:La Penséée Sauvage.

Gras, R., & Bailleul, M. (Eds.). (2000). *La fouille dans les données par la méthode d'analyse statistique implicative*.Caen: ARDM.

Gras, R., Couturier, R., Guillet, F., & Spagnolo, F. (2005, 30 mai - 1er juin). *Extraction de règles en incertain par la méthode statistique implicative.* Paper presented at the 12èmes Rencontres de la Société Francophone de Classification, Montréal, Canada.

Hacking, I. (1965). *The logic of statistical inference*.Cambridge: Cambridge University.

Hacking, I. (1971). Jacques bernoulli's art of conjecturing. *British Journal of Philosophical Science, 22*, 209-229.

Hacking, I. (2002). *L'émergence de la probabilité*.Paris: Seuil.

Hacking, I., & Dufour, M. (2004). *L'ouverture au probable*.Paris: Armand Colin.
Jaynes, E. (1984). The intuitive inadequacy of classical statistics. *Epistemologia, 7*(43), 43-74.
Jaynes, E. (1995). *Probability theory: The logic of science* (2003 ed.).
St. Louis, U. S. A.: Washington University.
Jeffreys, H. (1939). *Theory of probability* (1960 ed.). Oxford: University Press.
Jeffreys, H., & Wrinch, D. (1919). On some aspects of the theory of probability. *Philosophical Magazine, 38*, 715-731.
Jordan, C. (1926). Sur la probabilité des épreuves répétées, le théorème de bernoulli et son inversion. *Bulletin de la S.M.F., 54*, 101-137.
Kapitan, T. (1992). Peirce and the autonomy of abductive reasoning. *Erkenntnis, 37*.
Keynes, J. M. (1921). *A treatise on probability*.London: MacMillan and Co.
Kuzniak, A. (2004). *Paradigmes et espaces de travail géométriques.* Denis Diderot, Paris.
Laplace, P. S. (1771). Mémoire sur la probabilité des causes par les événements. *Mémoires de l'Academie royale des sciences de Paris, VI*.
Laplace, P. S. (1795). *Essai philosophique sur les probabilités* (1816 ed.). Paris: Pour les mathématiques et la Marine.
Lecoutre, B. (2005). Et si vous étiez un bayésien qui s'ignore? *Modulad, 1*(32), 92-105.
Leibniz, G. W. (1765). *New essays concerning human understanding* (A. Langley, Trans. 1916 ed. Vol. 1). Chicago: The Open Court Publishing Company.
Lerman, I. C. (1981). *Les bases de la classification hiérarchique*.Paris: Gauthier-Villars.
Neyman, J. (1977). Frequentist probability and frequentist statistics. *Synthese, 36*, 97-131.
Nicole, P., & Arnaud, A. (1662). *La logique de port royal* (A. Fouillée, Trans. 1878 ed.). Paris: Librairie Classique d'Eugène Belin.
Pascal, B. (1670). *Pensées* (É. Périer, Trans. 1995 ed.): Num. BNF de l'éd. de: Cambridge (Mass.).
Pearson, K. (1900). *The grammar of science* (Second ed.). London: Adam and Charles Black.
Popper, K. R. (1959). *The logic of scientific discovery*.London: Hutchinson.
Popper, K. R. (1959). The propensity interpretation of probability. *The British Journal for the Philosophy of Science, X*(37), 25-42.
Régnier, J.-C., & Oriol, J.-C. (2001). *Fonctionnement didactique de la simulation en statistique.* Paper presented at the Journées de Statistique Lyon 2003, Lyon, France.
Robert, & Collins. (2001). *Le robert*:VUEF.
Robert, A. (2001). Les recherches sur les pratiques des enseignants et les contraintes de l'exercice du métier d'enseignant. *Recherche en didactiques des mathématiques, 21*(1.2), 57-80.
Robert, A., & Robinet, J. (1993). Prise en compte du meta en didactique des mathématiques. *Cahier de Didirem, 21*.
Robert, A., & Rogalski, M. (2002). Comment peuvent varier les activités mathématique des élèves sur des exercices? *Petit x, 60*.
Shafer, G. (1996). *The art of causal conjecture*:MIT Press.
Shafer, G. (1996). The significance of jacob bernoulli's ars conjectandi for the philosophy of probability today. *Journal of Econometrics, 75*, 15-32.
Todhunter, M. A. (1865). *A history of the mathematical theory of probability.From time of pascal to laplace*.London: MacMillan and Co.
Vergnaud, G. (1990). La théorie des champs conceptuels. *Recherche en didactiques des mathématiques, 10*(2.3), 133-170.
von Mises, L. (1966). *Human action: A treatise on economics* (4° ed.). Chicago: Contemporary Books.
von Mises, R. (1928). *Probability, statistics and truth* (1981 ed.). New York: Dover.

Annexes

Annexe I

Programme et Documents d’accompagnement

Première ES et S

Annexe 2

Mathématiques

Série économique et sociale

Première - enseignement obligatoire

Nouveau programme applicable à compter de l'année scolaire 2001-2002

1 - OBJECTIFS GÉNÉRAUX POUR LA SÉRIE ÉCONOMIQUE ET SOCIALE

La science est un moyen ("déraisonnablement efficace" pour paraphraser Wigner) de rendre le monde qui nous entoure intelligible et partiellement prévisible: l'institution scolaire se doit de favoriser l'accès à ce moyen pour tous les lycéens, en particulier ceux de la série économique et sociale. Dans cette perspective, est réaffirmé ici le caractère indispensable d'un enseignement de mathématiques consistant dans cette série, et ce d'autant plus que par le biais des progrès technologiques, les mathématiques sont de plus en plus massivement présentes. Cet enseignement doit en particulier aider les élèves à intégrer des mathématiques dans leur mode de pensée; c'est là un travail de longue haleine et, à l'issue du cycle première-terminale, les élèves devraient avoir rencontré quelques types de questions appelant un traitement mathématique et saisi la nature des réponses que les mathématiques leur apportent.
Dans un premier temps les objectifs suivants seront prioritairement visés:
- entraîner à la lecture active de l'information, à sa critique éclairée et à son traitement, en particulier en privilégiant les connaissances et les méthodes permettant des changements de registre (graphique, numérique, algébrique,…);
- initier les élèves à la pratique d'une démarche scientifique globale, mêlant observation, exercice de l'imagination, questionnement, synthèse, usage de la logique, argumentation et démonstration mathématique;
- favoriser le travail personnel des élèves et donner la possibilité et le goût des problèmes consistants ou non entièrement balisés, qu'ils viennent des mathématiques ou d'ailleurs;
- promouvoir la cohérence de la formation des élèves en s'appuyant sur l'intuition, en relevant systématiquement les liens entre les différentes parties du programme et en exploitant les jonctions entre les mathématiques et les autres disciplines.

2 - MATHÉMATIQUES ET INFORMATIQUE EN PREMIÈRE ET TERMINALE ES

On peut souligner deux aspects du lien entre mathématiques et informatique.
- Utiliser des outils logiciels (sur calculatrice ou ordinateur) requiert des connaissances et des compétences mathématiques que cette utilisation contribue en retour à développer. Le programme insiste pour que cet aspect du lien entre mathématique et informatique soit travaillé à tous les niveaux; il ne s'agit pas d'apprendre à devenir expert dans l'utilisation de tel ou tel logiciel, mais de connaître la nature des questions susceptibles d'être illustrées ou résolues grâce à l'ordinateur et de savoir comment analyser les réponses fournies; l'élève doit apprendre à situer et intégrer l'usage des outils informatiques dans une démarche scientifique.
- L'informatique a totalement transformé le paysage des mathématiques; elle permet la confrontation aisée de plusieurs modèles, le calcul effectif de solutions non explicitables d'équations, la pratique de la simulation; des logiciels mettent à la portée d'un nombre toujours plus grand d'individus des applications de mathématiques sophistiquées, en particulier dans les entreprises. Une évolution des méthodes d'enseignement voire des contenus se fera peu à peu; s'il est nécessaire de l'amorcer dès aujourd'hui, il convient aussi de réfléchir et d'expérimenter diverses stratégies éducatives.
Le programme ne fixe pas de répartition entre différentes modalités qui doivent toutes être présentes: activités des élèves sur ordinateur ou sur calculatrice programmable graphique, travail de la classe entière (ou d'un groupe) utilisant un ordinateur muni d'un dispositif de visualisation collective. Il convient en ce domaine que les professeurs déterminent en chaque circonstance la stratégie d'utilisation la plus adaptée.

3 - ORGANISATION DE L'ENSEIGNEMENT ET DU TRAVAIL DES ÉLÈVES

L'horaire hebdomadaire est, pour la partie obligatoire, de 3 heures en première dont une demi-heure dédoublée et de 4 heures en terminale; s'y ajoutent 2 heures d'enseignement au choix en première et 2 heures d'enseignement de spécialité en terminale. Une cohérence forte s'impose entre les parties obligatoire et au choix (ou de spécialité); seule leur attribution à un même enseignant pourra réellement la garantir.
Chaque professeur garde toute liberté pour l'organisation de son enseignement, dans le respect des contenus et modalités de mise en œuvre précisés dans les tableaux du paragraphe suivant.
L'enseignant veillera à équilibrer les divers temps de l'activité mathématique dans sa classe; travail sur problèmes et exercices, élaboration de démonstration, exposé magistral, synthèse,… rythmeront les heures de classe et viseront tous à promouvoir chez chaque élève une

Annexe 2

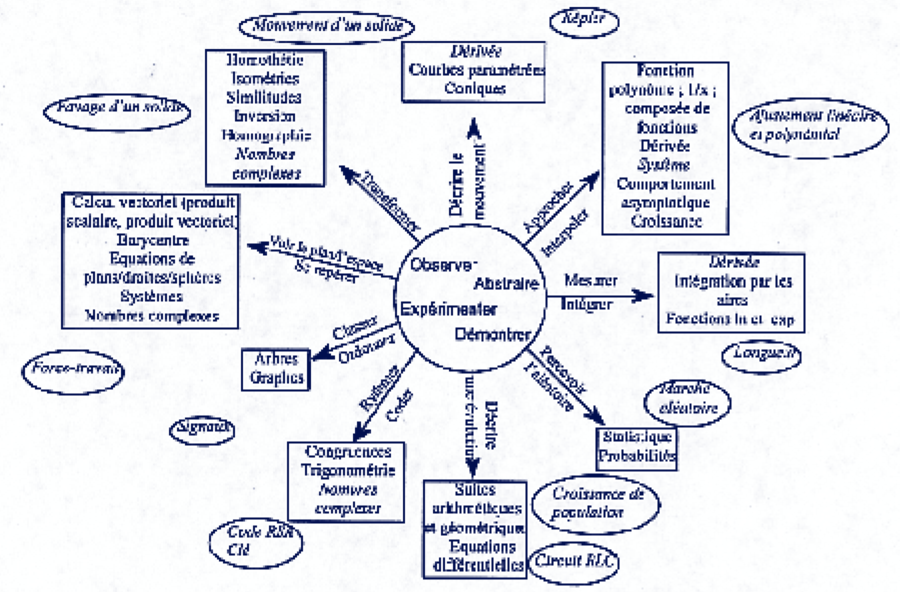
Mathématiques
Série scientifique
Enseignement obligatoire

Nouveau programme applicable à compter de l'année scolaire 2001-2002

1 - Généralités à propos d'une formation scientifique en première et en terminale S

Pour concevoir un programme de mathématiques dans le cadre d'une formation scientifique pour les élèves de première et terminale S, il convient :
- de prendre en compte la diversité des mathématiques actuelles;
- de rappeler les éléments fondamentaux propres à toute démarche mathématique et, de ce fait, incontournables dans tout projet de formation mathématique.

Le schéma suivant illustre ce propos; il permet par ailleurs de situer les choix de contenus définis au paragraphe 5.



•Le noyau central du schéma résume, en quatre composantes essentielles, la spécificité de toute pratique mathématique: observation, abstraction, expérimentation, démonstration. Ces quatre composantes entretiennent entre elles des rapports dialectiques, l'une appelant l'autre ou s'appuyant sur elle, au gré du travail mathématique réalisé.

Dans tous les domaines, l'observation est un processus dynamique suscité par une problématique propre à la discipline; elle conduit à des questions et éclaire ainsi l'origine et le développement de certaines idées. L'observation ne peut être pratiquée sans disposer d'un bagage théorique ; elle est d'autant plus riche que les connaissances de l'observateur sont importantes et organisées en un système cohérent. L'observation active demande de l'expérience et concourt en retour à la forger.

Annexe II

Transcription séances

Transcription séance Jeu de pièces de monnaie

Date : 13 Novembre 2006
N° élèves : 22

Observations :

- Premier colonne : temps [min]
- Enseignant : texte en *italique*
- Elèves : texte en **gras**
- Message adressée à toute la classe : colonne unique
- Message adressée à un ou plusieurs membres de la classe : deux colonnes

3 *Le travail que vous allez faire le travail qu'on va faire ensemble aujourd'hui bon çà y est au cour de la séance vous seriez vous feriez il y aura de moment où vous auriez accès à des ordinateurs pour pouvoir rentre certaines choses que vous auriez fait sur le sujet et voilà alors de quoi est la question et bien on a un jeu on va partir de çà le jeu c'est le suivant vous disposez chacun de quatre pièces de monnaie donc je vous donnerai quatre pièces de monnaie il va falloir jouer pour véritablement comprendre le jeu c'est le suivant bah le but du jeu*

4 *Le but du jeu c'est d'arriver on joue tout seule le plus vite possible à accumuler cinquante points alors comment est-ce qu'on accumule de points comment est-ce qu'on joue la règles c'est la suivante c'est que vous choisissez un nombre de pièces de monnaie d'une pièce à quatre et voilà vous dites je joue avec une pièce deux pièces trois pièce quatre pièces une fois que vous avez décidé le nombre de pièces de monnaie avec lequel vous allez jouer vous lancez ces pièces vous les lancez si une de ses faces si au moins une des faces est face donc si au moins une des faces supérieures est face vous ne gagnez rien et si vous avez sur les deux faces sur toutes le faces les faces supérieures de toutes vos pièces vous avez pile vous obtenez vous avez pile vous gagnez dix points*

5 *Donc çà quand vous jouez une fois que vous avez décidé le nombre de pièces de monnaie et bien vous jouez comme çà et le but est d'accumuler cinquante points bon eh combien j'en prend je ne peux prendre plus de cinq je peux en prendre quatre voilà j'en prend quatre je joue une fois là alors et j'ai obtenu pile face face face je je ne gagne rien j'en reprend les pièces je joue*

Quel côté est face et quel côté est pile on prend le côté que çà nous arrange quoi

Bonne question alors sur ces pièces la vous avez le un c'est quoi pile et là vous avez dessiné quoi là c'est qui d'ailleurs vous avez un visage là c'est face çà doit être l'Europe

6 *Donc on est bien d'accord pile c'est un bah c'est là où vous avez le numéro et face c'est l'autre d'accord bon donc il s'agit le but c'est de jouer et le problème qu'on va se poser c'est le suivant c'est puis qu'on a le choix au départ du nombre de pièce qu'on prend avec combien de le but est donc d'arriver le plus vite possible à accumuler les cinquante points donc avec combien de pièces est-ce qu'il convient de jouer*

Çà n'en change rien

Alors je vous propose je vous donne des feuilles de manière à que vous puissiez vous

pourriez rechercher je vous demande une chose c'est purement pour nous ces feuilles quelque soit ce que vous écrivez dessus prenez là comme un feuille de recherche
7 *je vous demanderez d'y mettre au moins votre prénom dessus et puis nous on ramassera c'est pour essayer de comprendre comment vous vous en servez comment vous çà vous semble et je vous donne des pièces quatre pièces à chacun*
8 **Au moins une des pièces supérieurs est face je ne comprend pas très bien**
Que si tu as plus de côté face que de côté pile
attendez et je vous propose aussi vous avez forcement des idées par exemple lui il a une idée une stratégie gardez-là pour vous pour le moment pour vous pour en suite débattre pour si tout le monde est d'accord et pour faire évoluer les idées
Réflexion de Florian vous m'avez des faces monsieur
Quand il y a égalité c'est quoi
En dix lancers j'ai eu cinq faces
Non, si tu as une face tu as pas gagné si pour deux pièces tu as une face qu'est pile tu as pas des points si les deux sont face tu n'as pas de points
Je n'ai que de faces moi
9 *Bon je vous laisse quelques instants pour comprendre le jeu pour prendre vous avez des idées vous essayez de les formaliser et après on en débattra*
Avec une pièce pas la peine de jouer
Ouais ouais tu as raison
En dix lancers j'ai cinq faces
Nombre de faces
Ah tiens regarde Dénis regarde
Quarante
Ouais
Et çà c'est quoi pourquoi tu fais çà
C'est le nombre de pièces que je joue
Et là
Nombre de piles de points quoi que j'obtienne
Tu es malade
Ouais çà c'est la classe
Là j'ai perdu
Ah jouer avec une pièce c'est plus sympa quatre pièces c'est impossible à gagner
12 *Avant que vous donnez vos avis vous les gardez pour quoi vous avez cette idée bon maintenant vous avez commencé là vous avez-vous pouvez arrêter les pièces quelques instants s'il te plaît maintenant que vous avez commencé vous avez commencé à comprendre le jeu à avoir des idées d'abord est-ce qu'il y a des questions sur la règle est-ce que tout le monde a compris la règle est-ce qu'il y a de points que vous paraissent obscurs sur la règle du jeu il y a en particulier une question qui m'a été posé là tout à l'heure j'insiste imaginez que je joue avec quatre pièces ou que vous jouez avec trois pièces si vous jouez avec trois pièces et que vous obtenez trois piles c'est chaque pile que vous donne dix points donc vous avez ne tout trente points çà été une remarque qui m'a été faite bah une remarque c'était une question*
13 *Est-ce que vous avez d'autres questions bon est-ce que vous avez compris le jeu là que*

vous comprenez bien les règles qu'est-ce que vous avez comme idées de ce jeu
Bah que plus on prend de pièces moins on a de chances d'avoir que de piles
Plus tu prends de pièces
Moins de chances d'avoir que de piles
Bah pas trop
Par contre si on gagne on gagne plus
Ouais
Que par contre si on gagne on gagne plus
Une chance sur deux
Par contre si on gagne on gagne plus
Oui mais il faut plus vite
Bon on va essayer de s'écouter chacun peut parler bien sur mais on s'écoute
Tu gagnes plus vite
Tu gagnes plus vite
Si on joue avec quatre pièces
On gagne pas forcement plus vite
Attendez attendez attendez chacun à son tour
On gagne pas forcement plus vite autant il faut voir qu'on mets plus le temps pour avoir quatre piles
14 **En un lancer avec deux pièces lancées au même temps tu fais vingt points**
Si toutes sont pile quoi
Moi j'ai essayé j'ai pas réussi beaucoup de fois
Si on joue avec quatre pièces moins de chances de gagner
En un lancer
Moins il y a des pièces plus on a des chances de gagner
Oui vas-y toi
Moins il y a des pièces plus de chances de gagner vite
Vite
En général
C'est général çà
Oui, oui en général je dis
Quelqu'un veut réagir à ce qu'il vient de dire oui vas-y plus il y a de pièces
Plus il y a des pièces plus la chance de réussir est faible
15 *Dis-le plus il y a des pièces plus la chance de réussir est faible il y en a qui me disent oui j'entendu aussi d'autres choses*
En moins de lancers quoi
Mais il y aura moins des lancers
Il y aura moins des lancers si on prend quatre pièces
Non
Non çà a rien à voir
Mais non çà a rien à voir
Çà a rien à voir écoutez que c'est intéressant tout çà
Bah si on joue avec une pièce on a plus des chances de tomber sur pile si on a beaucoup de pièces on a moins des chances de tomber sur de pas tomber sur face plutôt
Ouais
Bah le nombre le nombre de points récoltés avec un grand nombre de pièces est grand mais plus rare aussi
Plus rare ouais c'est vrai c'est intéressant ce que tu dis
16 *D'autres ont des avis convergentes complémentaires des choses que vous n'êtes pas*

d'accord bon alors ce que je vous propose ce que je vous propose maintenant c'est que pour rentrer quand même dans le jeu pour rentrer dans le jeu vous allez maintenant chacun rapidement faire à votre tour simplement un lancer un lancer avec une pièce un lancer avec deux pièces un suite un lancer avec trois pièces en suite un lancer avec quatre pièces et puis notez ce que vous obtenez et puis on va regarder on va regarder ce que cela donne d'accord donc vous le notez bien vous le notez bien et puis je le noterez

Il y a une façon de lancer si il y a une façon de lancer

Que s'il a une autre façon de lancer

Est-ce que il y une façon de lancer

Qu'est-ce que tu veux dire par çà

Bah comment je fais comme çà ou

Ah tu veux dire

Parce que moi je le lance comme çà à plat

Oui tu le lances comme çà à plat et quand on lance une pièce de monnaie quand on joue au hasard

Ah d'accord

On indique le nombre de points qu'on a réalisé

Ouais tu indiques tu indiques le nombre de points que tu réalises le nombre de points que tu réalises pour une pièce pour deux pièces pour trois pièces comme çà

18 *Alors pour un tu as eu combien*

Toi

A chaque fois le un arrive souvent

On continue

Dix zéro zéro zéro

19

Zéro par tout

20

21 **Dix zéro trente zéro**

Maintenant qu'on a fait ce tableau est-ce qu'on a nous allons nous en servir chacun pour faire pour observer si vous observez quelque chose qu'est-ce que vous observez

22 **Bah qu'il y a un seule qu'a réussi à avoir quarante le reste**

Que les dix çà va vite

Que les dix çà va mais le reste

Une chance sur deux et une chance sur quatre

Comme on a dit tout à l'heure plus il y a des pièces

Attends comment tu dis

Que plus il y a des pièces moins on a des chances de réussir

Pas la grosse chance

Il faut que le jeu soit équitable

On continu

C'est pas équitable du tout pas du tout il faut voir les points

Il suffit de faire il suffit de faire la somme de chaque colonne quoi et voir le nombre de points qu'on a eu par colonne quoi

Bon fessons çà rapidement il nous suggère de faire la somme donc la somme pour quatre c'est facile ici quarante donc ici la somme c'est quarante je ferai un deux trois quatre donc la somme c'est quarante

23 **Cent cinquante la première**
Cent cinquante à la première et à la trois
Cent cinquante à la première à la trois aussi
Et quatre vingt au deux
quatre vingt au deux
Bon alors vous en pensez quoi
Un et trois c'est égal au peu près
Non ce qu'il faut savoir
Pour moi c'est pareil
Attendez attendez attendez chacun à son tour chacun a son idée on s'écoute et puis après vous réagissez alors toi tu disais
Un et trois c'est le même c'est
Tu constates là qu'un et trois c'est le même est-ce que çà donne envie de conclure et de répondre à la question sur la stratégie qu'il faudrait avoir pour gagner le plus rapidement possible et arriver le plus rapidement possible à cinquante points
Mains non parce qu'il faudra faire plusieurs séries comme çà çà peut pas être un cas général çà
Çà peut pas être un cas général c'est vrai que bah
24 *Ici on a fait on n'a fait que vingt deux vous avez fait chacun un tirage est si on fait on refait la même expérience est-ce qu'on va obtenir la même chose*
Non non
Non
Non
Non
Est-ce qu'on va obtenir la même chose est-ce qu'on va obtenir la même chose est-ce qu'on va obtenir quelque chose de contraire
Bah non on ne sait pas
On ne sait pas
On ne peut pas parce qu'il y a du hasard
C'est du hasard donc on ne peut pas savoir
Mais il y a quand même
Mais c'est une approximation
Il y a quand même une probabilité
Attendez attendez
Il y a quand même une probabilité que l'un reste dans les alentours de cent cinquante bah pas exactement cent cinquante mais
Bah on fait l'expérience si vous voulez
Ouais
Ouais
C'est çà que je disais
Donc c'est parti
25 **…**
26 **Zéro vingt zéro zéro**
27 **…**
On a plus de quarante déjà
28 Bon et çà fait combien là
Quatre vingt dix
Cent quarante et le trois
29 **Et deux cent**
Bon comparons quand même on compare avec ce qu'on avait trouvé à l'autre dans les

faits

C'est pas pareil

En effet, ce n'est pas pareil est-ce que çà nous permet de est-ce que çà nous permet de mieux comprendre le jeu est-ce que çà confirme ou au contraire çà infirme

Il y a du hasard

les idées que vous aviez sur le jeu alors vos réactions on entend c'est du hasard qu'est-ce que tu veux par hasard

30 **Qu'on ne sait pas du tout comment va sortir pour une pièce pour deux pièces pour trois pièces on ne peut pas savoir**

C'est égal qu'on lance une pièce quatre fois ou quatre pièces une fois c'est égal pour moi il y a autant de chances de gagner

Et qu'est-ce que permet de penser çà

Pour moi c'est égal il y a autant des chances de gagner

Qu'est-ce que te permet de dire çà

Bah les données

Les données tu vois çà tu vois l'autre tiens pourquoi ton idée en voyant çà tu dis il y a autant de chances de gagner on va la stratégie elle va être la même en jouant avec une deux trois ou quatre pièces

Bah oui parce qu'on lance autant les pièces on gagne moins bah avec une pièce on gagne moins pour moi c'est pareil

C'est pareil est-ce que çà apparaît dans les données est-ce que pourquoi c'est simplement que tu expliques je dis pas que tu as tort ou que tu as raison qu'est-ce que dans les données qu'on vient de faire

qu'est-ce que dans les calculs qu'on vient de faire te fait dire

Bah auprès le premier tableau celui qui gagne le plus c'est une seule pièce

Et trois pièces

Et trois pièces aussi tandis que dans le deuxième tableau c'est plus le cas c'est différent c'est quatre c'est deux et quatre

Ouais c'est quatre bon est-ce que c'est l'idée qu'avait est-ce que c'est cette idée que tu avais au départ

Non

Quand tu as présenté c'était quoi ton idée au départ

Moi je pensais qu'on gagnait plus avec une seule pièce

Qu'on gagnait plus avec une seule pièce et là tu te convaincre du fait que finalement çà revient au même d'autres

32 *D'autres idées d'autres avis tu as une remarque à faire*

Bah c'est juste une analyse quand on prend une seule pièce on atteint les cinquante

C'est çà que je voulais dire

En peu de coups mais quand on joue avec quatre pièces on attend le cinquante en je sais pas combien de lancers bah le but c'est d'arriver à cinquante

Mais si on regarde le tableau lequel qui est arrivé le premier à cinquante à chaque fois

On a une chance sur deux

Bah c'est un ou deux

Çà dépend

Oui mais un ou deux quand même

Une chance sur deux

Je ne comprends pas ton raisonnement

En partant d'en haut en partant d'en haut pour arriver à cinquante le plus vite

Vas-y termine ton raisonnement attend attend attend pour arriver à cinquante le plus vite

si tu regarde ici tu as un deux trois quatre cinq si c'était le même jouer bah si c'était le même jouer qu'avait joué en fait c'est pas le même jouer j'en insiste mais en effet
Il aurait gagné l'un
Il aurait gagné avant
ici on serait arrivé rapidement dans ce cas c'est vrai qu'on aurait arrivé largement

33 *Ouais c'est est-ce que vous trouvez regardez si je commence si on imagine que çà c'est le début de la partie ici c'est me début d'une partie celui qu'a choisi en imaginant que le jouer donc il est qui commence et il tire une fois trente il tire une fois trente çà veut dire on a gagné et il a gagné plus rapidement que celui-là donc là c'est non plus bah cette lecture du tableau est-ce que vous pouvez trouver d'autres bah d'autres jeux à la lecture de ce tableau quand on joue un on a tout le temps bah on va plutôt gagner bah d'après le tableau moi je t'ai trouvé un contre exemple en imaginant je joue çà est-ce que vous voyez d'autres*
Vingt et un et vingt deux

34 *Vingt et un et vingt deux bah c'est vrai que vingt et un et vingt deux c'est quarante le vingt et un et vingt deux*
Non trente et vingt
Voilà par exemple le vingt c'est impossible c'est impossible d'avoir trois vingt quasiment d'affilé
C'est pas possible d'avoir trois vingt d'affilé
C'est beaucoup plus dur quoi
Si on regarde el tableau en bas on a deux fois quarante de suite
Mais par rapport au nombre de faces il va avoir une chance sur deux de gagner d'avoir dix
Ouais
Ouais
De gagner et d'avoir dix points
Il faut voir les pourcentages
Bon on va rester avec une face donc tu parles d'une face
Bah soit on tombe sur face soit on tombe sur pile si on tombe sur pile on gagne dix points si on tombe sur face on va rien gagner
Ouais alors est-ce que çà çà te permet de

35 *çà te permet de décider de juger de caractériser le jeu quand on prend une pièce*
En dix lancers normalement on arrive à cinquante
Ouais en dix lancers
En dix lancers
On a une chance sur deux de tomber sur pile en dix lancers normalement on devrait arriver à cinquante
Sur dix lancers tu peux toujours tomber sur face
Oui en moyenne en moyenne
En moyenne
En moyenne
C'est une moyenne quoi
Alors est-ce que vous entendez là on parle de moyenne vous parlez de moyenne qu'est-ce que vous entendez par moyenne ou est-ce que vous pouvez le reformuler là pour tout le monde
Bah qu'on a une chance sur deux que çà tombe sur pile
Il y a une chance sur deux que çà tombe sur pile çà vous êtes tous d'accord il y a une chance sur deux que çà tombe sur pile et là et après vous me parlez de moyenne est-ce que vous pourriez le préciser pour tout le monde

36 **Qu'il y a une chance sur deux quand on lance la pièce mais au bout de dix coups pour arriver à quarante points on ne peut pas savoir**
On peut pas savoir

Une chance sur deux par coup

Non mais qu'on ait cinquante points accumulés

Mais si la pièce elle avais tombée sur pile il faudrait aller jusqu'à vingt fois
Oui tu as une chance sur deux que lors de dix coups tu atteignes les cinquante points
Là il faut être très chanceux
Donc là il y a deux points de vue il y a deux points de vue là qui sont deux points de vue intéressants là vous vous êtes vous êtes au cœur du problème redonne bien toi ton point de vue
Moi c'est en fait que on a une chance sur deux de tomber sur face mais sur dix coups bah sur pile mais pour cinq coups d'affilé on ne peut pas savoir bah je ne sais pas
Le hasard
37 *Donc on ne peut pas savoir dans le sens on ne peut pas prédire maintenant on ne peut pas prédire*
Bah non
Maintenant on ne peut pas prédire oui c'est vrai je t'ai coupé
Pour deux coups deux lancers c'est une chance sur quatre pour trois lancers c'est une chance sur huit
Une chance sur six
Neuf
Sur neuf
Sur neuf et pour quatre pièces c'est une chance sur seize
Mais après sur dix lancers on peu pas savoir
Saisissez vous le débat qui s'instaure
A chaque fois il y a deux faces donc
Attends tu le redonne vas-y
A chaque fois il y a deux faces donc il y a deux possibilités par pièce
On travaille sur une pièce après on fait pour les autres après on généralise pour deux pièces pour le moment une seule pièce
38 *Pour le moment ce que j'entend il y a une idée quand il y a une seule pièce de toute façon il y a une chance sur deux d'obtenir pile on a une chance sur deux d'obtenir face et il y a une autre idée qui est venu c'est qu'on a parlé de moyenne qui a parlé de moyenne Dennis qu'est-ce que tu voulais dire quand tu parlais de moyenne*
Non non bah c'est çà ce que je voulais dire c'est que je me suis mal exprimé quoi
Et c'est quoi ce que tu voulais dire
Bah une chance sur deux
Sur dix lancers
Ouais ouais sur dix lancers quoi
Alors tu m'as parlé de dix lancers tu m'as parlé de dix lancers pourquoi tu m'as parlé de dix lancers
Bah parce que c'était au début sans réfléchir
Comme il y a une chance sur deux
au début je voyais une chance sur deux pour un lancer mais pour dix lancers je sais pas c'est pas une logique
Absolue
au but de dix lancers on n'a pas une chance sur deux d'accumuler cinquante points
Bah si

Bah non
39 **Non mais il y a une chance quand même**
Oui mais elle
Elle n'est pas absolue quoi
Mais sur dix lancers de pièces
Il y aura cinquante points si tu as dix lancers bah mathématiquement après c'est autre chose
Mais c'est pas sur
Bah si si si sur un nombre infini de lancers
Non tu n'as aucune
C'est comme la chance d'avoir une fille ou un garçon c'est une chance sur deux
Mais non la preuve c'est qu'il y a plus de gars que de filles et ce n'est pas une chance sur deux
Çà a toujours été je sais pas
40 *Attends redonne bah continue avec ton argumentation tu as d'autres choses à apporter bon revenons on va progresser on va progresser on va essayer de revenons à notre problème c'est de décider quel est le nombre de pièce qu'il faudrait prendre alors pour le moment j'ai entendu dans les dans les stratégies j'ai entendu deux arguments j'en ai entendu au début m'ont dit on va prendre une pièce on va prendre une pièce parce que çà a plus de chance çà a plus de chance gagner c'est ce que j'ai entendu j'ai entendu un petit peu la même chose dans le tableau ou sur tout en regardant*
41 *pas celui-là mais l'autre tableau vous me disiez attendez*
Non mais çà sur vingt lancers
Vous me disiez que qu'on voit dans ce tableau que quand on joue quand on joue avec une pièce on arrive assez souvent on arrive assez souvent à cinquante plus rapidement c'est ce que j'ai entendu et puis un autre argument que j'ai entendu Romain an autre argument c'est que finalement qu'on joue avec un deux trois ou quatre pièces de monnaie çà va revenir au même et pourquoi çà va revenir au même ceux qui pensaient là
Parce qu'on a la même chance de tomber sur bah les quatre piles et si on tombe dessus
42 *Qu'est-ce qui fait la difficulté de ce jeu qu'est-ce qui fait la difficulté pour se prononcer sur la stratégie qu'est-ce que selon vous fait que vous avez du mal à vous prononcer*
Bah la chance
C'est le hasard
Çà dépend qui lance la pièce
Le cinquante
C'est que le score à attendre c'est cinquante
C'est le hasard c'est que le score à attendre c'est cinquante c'est-à-dire
Il faut qu'on ait deux fois quarante
Monsieur là c'est cent cinquante sur vingt deux lancers
Sur vingt deux lancers on a obtenu cent cinquante
Oui mais de l'autre côté c'est pas pareil
C'est pas du tout pareil c'est pas du tout pareil
On ne peut pas se prononcer
On ne peut pas se prononcer est-ce qu'on pourrait se prononcer est-ce qu'il y a un moyen de se prononcer
Oui mais pas sur les tirages
Il faudrait faire plus de tirages
Ouais beaucoup plus de tirages
Une infinité de tirages pour être sur

Il faudrait faire beaucoup plus de tirages
La probabilité
Pour attendre cinquante
Une infinité
Mais çà n' change rien
Pour attendre cinquante il faut deux fois trois et deux fois quatre et il est plus dur d'avoir deux fois quatre due d'avoir deux fois trois
Oui parce qu'il moins de pourcentage de chances de tomber sur quatre pièces que sur trois donc déjà oui
Si tu tombes sur quatre pièces tu gagnes quarante et avec trois pièces tu as trente
Oui mais on n'atteint aussi facilement cinquante il faut l'avoir deux fois
Mais il faut l'avoir deux fois quand même
Oui mais çà pour les deux
Ouais mais n'on attends aussi facilement quarante et il faut l'avoir deux fois
Et il avoir deux fois je comprends ton point de vue tu veux dire si je je le dis autrement tu me dis si çà correspond à ton idée tu veux que quand on joue avec trois ou avec quatre quand on gagne on gagne plus on est forcement au-delà de cinquante
Ouais
C'est çà c'est un peu çà ton idée donc on gagne plus qu'était dans la règle au début quoi
44 *Et donc c'est moins intéressant*
Non c'est que c'est plus facile d'avoir deux fois trois que d'avoir deux fois quatre
Oui c'est çà il y a plus de chances
Alors il y a un mots qui a été aussi prononcé il y a un mot qui a été prononcé c'est l'idée de moyenne et c'est une question qu'on va se poser c'est en moyenne combien est-ce qu'on gagne en moyenne est-ce qu'on est capable d'estimer combien on gagne si on joue avec une pièce si on joue avec deux pièces si on joue avec trois et quatre pièces
Ouais la moyenne on peut la faire mais c'est par autant que même si par exemple sur la pièce de jeux on a une bonne moyenne par exemple par rapport à la chance c'est par pour autant
Oui mais c'est théoriquement çà
Ouais mais c'est pas pour autant que même si il a par exemple on a vingt quarante pour cent de chances de gagner avec une pièce de vingt c'est pas pour autant qu'on va réussir quand on va le faire çà dépends après
Oui mais c'est pour cela que c'est une moyenne
Oui mais là tu peux pas te fier sur laquelle tu ne peux pas te fier pour faire tes lancers
Oui oui bien sur c'est une moyenne
45 **A gagner**
Ouais mais c'est pas pour autant qu'on va réussir
Bah ouais
Bon on a aussi dit qu'on est limité par la taille du tableau c'est ce que vous m'avez dit on ne peut pas se cantonner à ces simples résultats pour pour se décider bon alors pour progresser d'abord mettons nous bien d'accord sur ce qui nous permet de sur ce qui nous permet de décider nous avons une stratégie un but d'arriver d'arriver le plus vite possible le minimum de coups à cinquante points pour arriver à cinquante points avec un minimum de coups
46 *Et bien une manière de se prononcer çà serait de se dire de décider qu'on essaie de gagner en moyenne de gagner en moyenne le maximum de points par tirage d'accord et puis si on arrive à trouver çà et bien on arrivera à se prononcer sur plutôt en choisissant le nombre de pièces en fonction de tout çà on va choisir le nombre de pièces que permet*

de gagner le maximum de points par tirage d'accord bon alors cette idée l'idée maintenant que va nous intéresser c'est cette idée de gagner en moyenne gagner ne moyenne le maximum de points autrement dit

47 *Je prononce autre mots je vais parler de tendance et on voudrais connaître la tendance de ce jeu qu'est-ce qui se passe des idées qu'on été dites tu nous as parlé de tout à l'heure tu as prononcé dans cette idée de tendance qu'est-ce que tu as prononcé tu as parlé de moyenne mais tu parlais de faire une infinité de fois tu as prononcé le mots le faire une infinité de fois bon avant de faire çà avant de avant on va s'intéresser donc on va prendre pour chacune des pièces pour chacune des nombres pour chacune des possibilités et alors on va essayer de regarder pour chacune de ces pour chacun de ces cas on va regarder combien on va regarder combien en moyenne*

48 *on obtient de points par tirage et puis en fessant çà on aura peut être à se prononcer pour une stratégie la meilleure bon alors on prend une pièce on prend une pièce j'ai écris j'ai déjà écris ce que j'ai entendu d'une manière ou d'une autre si on joue avec une pièce de monnaie ce qui est vrai ce que parfois on gagner parfois on gagne pas là c'est sur mais en moyenne en moyenne on gagne combien*

Une chance sur deux

Oui cinquante pour cent de chances

Ouais

Alors on a une chance sur deux de gagner je suis d'accord qu'on a une chance sur deux mais là je te parles

Soit on gagner zéro soit on gagne dix

Si je tranche c'est cinq

49 *Soit on gagner zéro soit on gagner dix tu me dis tu as une chance sur deux de gagner si je gagne si je gagne tu gagnes combien*

Bah dix

Dix points si je gagne je gagne combien

Si je perds rien

Si je perds je ne gagne rien donc qu'est-ce que tu peux gagner à chaque fois que tu joues on peut gagner dix mais en moyenne est-ce que tu crois qu'on va gagner dix tu vas gagner cinq tu me dis pourquoi tu vas gagner cinq on s'écoute s'il vous plait

Bah il y a une chance sur deux

Une chance sur deux

Une chance sur deux et comme on a une chance sur deux et comme on peut gagner dix bah finalement en moyenne on va dire tiens tu vois en moyenne on gagne cinq bon est-ce que c'est est-ce que c'est une tendance qu'on retrouve dans ce qu'on a fait c'est une tendance qu'on aura du mal à voir apparaître parce que on s'a limité à très peu de tirages

ici on en avait vingt deux on a vu qu'en moyenne sur vingt deux tirages vous en aviez quinze vous en aviez quinze maintenant vous en aviez quinze sur vingt deux donc ici là la moyenne de gains est de combien elle est de combien la moyenne de gains quand on a lancé une pièce là sur les vingt deux sur ces vingt deux lancers vingt deux cent cinquante cent cinquante sur vingt deux faites cent cinquante divisé par vingt deux çà fait combien

Çà fait six quatre vingt un

çà fait six quatre vingt un vous par votre raisonnement par votre raisonnement vous me parlez de probabilité d'un demi vous arrivez à cinq

51 *si je regarde l'autre tableau là je regarde l'autre tableau j'y arrive on était à combien là quatre vingt dix sur vingt deux*

Quatre et quelques ..

Bon donc voyez qu'ici dans ces vingt deux tirages dans ce qu'équivaut à faire vingt deux *tirages et bien en moyenne on a gagné quatre virgule quelque chose et puis ici on a gagné sept et quelques regardez grâce à ce grâce à ces deux séries de vingt deux tirages on pourrait faire comme si on aurai fait quarante quatre tirages et comme çà on aurait quarante quatre lancers sur les quarante quatre lancers on a gagné combien*

52 *Sur les quarante quatre lancers on a gagné cent cinquante plus quatre vingt dix çà fait deux cent quarante et bien çà fait combien donc en moyenne sur les quarante quatre lancers sur les quarante quatre lancers on a gagné combien*

Cinq virgule quarante cinq

On a gagné cinq virgule quarante cinq voyez qu'un coup on était à quatre l'autre fois on était à sept quand on fait vingt deux mais quand on s'intéresse d'un coup à plus de lancers et bien là on se rapproche de quelque chose et cette quelque chose vers la quelle on se rapproche c'est le cinq que vous avez réussi à sentir bon maintenant on va s'intéresser pour vraiment décider on va s'intéresser aux deux pièces de monnaie bon alors avant de bah avant est-ce que vous avez réussi à faire un raisonnement avec une pièce vous me dites qu'en moyenne on trouve cinq

53 *est-ce que vous pourriez me faire un raisonnement analogue est-ce que vous pourriez me faire un raisonnement analogue avec deux pièces de monnaie*

Cinq

Tu peux nous c'est ton raisonnement écoutez s'il vous plait

Tu ne peux pas le sortir comme çà sans rien expliquer

Bah on a quatre cas le premier quand la première de deux pièces tombe sur pile la deuxième sur face on gagne zéro deuxième la première tombe sur face la deuxième sur pile on gagne zéro la troisième les deux tombent sur face on gagne zéro et quatrième les deux tombent sur pile on gagne vingt points vingt divisés par quatre çà fait cinq

Et vingt divisé par quatre çà fait

Çà fait cinq

Donc vingt divisé par quatre çà fait

Cinq

Çà fait cinq vous avez compris ce raisonnement

Ouais

Ouais

Est-ce que quelqu'un a un autre raisonnement à nous proposer

54 *Un autre raisonnement à nous proposer tu as convaincu par son raisonnement pas trop bon si j'ai bien compris donc la moyenne de gains çà serait quand on prends deux pièces*

Cinq

Çà devrait être cinq est-ce que vous en êtes convaincus

Ouais

Visiblement il y en a qui me disent oui quand j'interroge Florian il dit non et je pense il n'est pas le seul il n'est pas le seul vous êtes pas complètement convaincus on fera pas de pause on va continuer non non sinon tu peux aller voir le règlement intérieur donc on continu là

55 *Puisqu'on a du mal est-ce que d'autres est-ce que dans la classe je le répète là je le répète pour être sur qu'il n'y a pas certains qui n'osent pas se prononcer qui auraient fait un autre raisonnement à celui-là et qui auraient trouvé une autre moyenne bon on va faire un petit sondage qui a compris le raisonnement de Jeremy et qui est d'accord avec le fait que la moyenne de gains avec deux pièces de monnaie c'est de que la moyenne de gain c'est cinq un deux quatre six huit dix douze quatorze seize qui n'est pas qui n'est pas convaincu personne*

56 **Bah je sais pas si je**

Bah si si tu n'est pas convaincu tu lèves la main

Pas convaincu moi

Tu n'est pas convaincu bon ce que je vous propose c'est qu'avant de revenir au raisonnement de de revenir sur le raisonnement qu'il a fait c'est de prendre un temps et se dire on va essayer on va essayer de faire des expériences comme tu le disais tout à l'heure comme tu l'as dis tout à l'heure pour essayer de se prononcer si on fait des lancers de deux pièces il faut en faire beaucoup donc on va en faire un certain nombre alors voilà comment on va procéder vous allez chacun vous allez chacun effectuer dix lancers de deux pièces Florian c'est fini vous notez les tirages les tirages obtenus bien sûr alors on va coder de la manière suivante si vous obtenez pile pile vous codez 1 là

57 *J'ai fait une espèce de petit tableau et en suite on va saisir chacun vous allez saisir c'est pour cela qu'on a des ordinateurs portables vous allez saisir et en suite on va tout rassembler donc pour un élève ah vous faites dix tirages vous notez un si vous avez deux piles et puis zéro si vous avez pas de piles une fois que vous auriez fait çà vous le notez vous aurez fait çà on va vous faire passer les ordinateurs portables et puis sur l'ordinateur portable vous allez rentrer vos vous rentrerez vos résultats*

58 *Dans le tableau qu'on vous fera passer dans le tableau qu'on vous fera passer il y a des regroupement par dix dans les colonnes qui sont en couleur vous rentrez vos valeurs*

…

71 *Çà y est vous avez eu votre pause arrêtez les pièces s'il vous plait donc ce que j'ai devant vous çà va être de rassembler tous les résultats alors dans le fichier qu'on a rassemblé les résultats vous allez voir au four et à mesure vous allez voir un graphique qui apparaît qu'il faudra commenter*

72 *J'aimerais je vous demande une chose c'est un peut çà devient un peut fatiguant là tout le monde est un peu fatigué concentrez vous sur ce que je vous montre donc voilà je vais copier là je vais recopier là-dedans les différentes tirages que vous avez fait alors je prends la première série de tirages*

Je dis pas que c'est forcement cinq la probabilité quand on est à vingt avec une pièce c'est cinquante et après c'est

73 *Donc j'ai fait on collage on va essayer de commenter on va essayer de comprendre est-ce que vous arrivez à voir don là vous avez les différents tirages que vous avez effectué là un élève qui a effectué ses dix tirages et ici il y a un calcul qui se fait le calcul qui se fait ici c'est un calcul de pourcentage*

74 *C'est un calcul de pourcentage si dans le premier tirage on a eu zéro pour un tirage ah il y a eu pile pile donc la fréquence d'apparition de pile pile est en moyenne on a obtenu zéro fois bon sur deux tirages on a cinquante pour cent parce que sur deux tirages on a eu un qui a donné pile pile et il y a cinquante pour cent d'apparition et au même temps vous voyez qu'il y a un graphique qui va se faire ici le graphique que va se faire ici et bien c'est par exemple quand on est au four et à mesure que je rentre les tirages on va voir apparaître l'évolution de ce calcul de pourcentage qu'on soit bien d'accord ici là le zéro trente quatre le zéro trente quatre de qui apparaît çà veut dire quoi*

75 *Comment est-ce que vous l'interprétez là je suis à la colonne je suis à la ligne soixante dix ligne soixante dix et j'obtiens zéro trente quatre il y en a qui m'énerve maintenant je reviens pour que vous participez au lieu de faire autre chose donc le répète ma question ce zéro trente quatre qu'apparaît qu'est-ce qui signifie oui*

C'est la probabilité qu'en soixante dix coups

Alors en soixante dix coups on va pas parler de probabilité on va parler de pourcentage et il y a eu trente quatre zéro trente quatre trente quatre virgule vingt huit pour cent de coups qu'on a obtenu pile pile d'accord donc vous avez en même temps le graphique et ici voyez que soixante soixante douze

75 *Lancers et bien on est à on a dit qu'on était à une fréquence bah on avait un pourcentage qui était de zéro virgule trente zéro virgule trente quatre d'accord bon j'ai rentré les autres j'ai fait la première la première série le premier groupe d'élèves je vais prendre la deuxième le deuxième groupe d'élèves*

76 *On regarde le graphique voyez je suis maintenant à cent vingt cent trente deux j'ai mis cent trente deux tirages on a dix tirages là on a un pourcentage qui est de l'ordre de zéro virgule vingt neuf donc sur le cent trente deux tirages j'en ai vingt neuf pour cent ah j'en ai vingt neuf pour cent qui ont donnés pile pile je prends l'autre le dernier*

78 *On fait le graphique alors qu'est-ce que vous en pensez*

est-ce que çà vous parait cohérent par rapport à

Bah oui

Par rapport à vos idées par rapport à ce qui par rapport à ce qui a été dit on se rapproche de zéro virgule vingt cinq

Un chance sur quatre

On se rapproche de zéro virgule vingt cinq bon alors est-ce que maintenant on peu revenir sur ce zéro virgule vingt cinq est-ce que quelqu'un peut venir au tableau et nous expliquer comment il intervient cette cette probabilité zéro virgule vingt cinq quelqu'un veux nous Jeremy ta manière de faire

79 *Tu veux venir nous expliquer on entend des choses qui personne pour expliquer vas-y mais tu n'as pris que dix tirages*

Ouais ouais pour deux pièces on a deux quatre possibilités on a deux faces par pièce on a deux faces par pièce et il y a quatre possibilités

Donc il y a deux faces par pièce

On a quatre possibilités de faire des tirages

81 **Je mettrais P pour pile et F pour face donc on a soit pile pile PP PF FP FF et on a un zéro zéro zéro**

Une chance sur quatre

Donc on a deux cent vingt tirages çà fait cinquante cinq

Alors çà veut dire quoi ton cinquante cinq tu peux l'interpréter

82 **Le nombre de possibilités qu'on va avoir de trouver deux fois d'avoir un sur le nombre de tirages fais**

C'est pas logique

Sur le nombre de tirages qu'on va faire alors il y a cinquante cinq lancers oui vas-y

Sur cinquante cinq lancers on a une chance d'avoir que des piles

Sur cinquante lancers on devrait avoir pile pile

Ouais à peu près

A peu près

On a une chance sur quatre

Cinquante cinq sur deux cent vingt

83 *Il faut pas que tu appuies sur le petit bouton*

Cinquante cinq sur deux cent vingt çà fait zéro virgule vingt cinq mais comment tu as trouvé le zéro virgule vingt cinq tu as fait cinquante cinq divisé par deux cent vingt

Ouais

Bah non

D'accord là c'est ta manière d'expliquer quelqu'un a autre manière va nous l'expliquer au tableau et comme çà on aura une trace avoir le calcul qu'on va faire parce que lui il a réussi à trouver il nous dit que la probabilité d'obtenir pile pile c'est zéro zéro virgule vingt cinq et il a fais ses calculs pour le trouver comment toi tu fais pour le faire

C'est le nombre de possibilités sur l'ensemble de possibilités

C'est-à-dire

C'est l'ensemble

Il y a combien de possibilités

Quatre

Il y a quatre possibilités

83 **Il faut appuyer plus fort**

Donc il y a quatre possibilités et tu dis que la probabilité c'est quoi

Un sur quatre

Vas-y écris-le un sur quatre bon alors qu'est-ce qui vous convainque vous êtres convaincus par les calculs qu'ils ont fait maintenant mon problème c'est de savoir je vous rappelle que dans notre idée pour essayer de déterminer pour chacun pour chaque possibilité une pièce deux pièces c'est en moyenne combien

84 *quel est le gain pour chaque tirage en moyenne qu'est-ce combien on va gagner*

Cinq

Donc ici qu'est-ce que tu fais pour déterminer ces cinq

Vingt fois zéro virgule cinq

C'est vingt fois un quart tu as une chance sur quatre tu as une chance sur quatre d'obtenir pile pile quand on obtient pile pile et bien tu gagnes vingt donc en moyenne tu fais vingt fois un quart et tu trouves çà bon d'accord bon alors essayons de clarifier le résultat et revenons un petit peu au vocabulaire qu'on emploie quand on fait de probabilité qui nous servie pour prouver et qui va nous aider à faire d'autres calculs

85 *Bon alors la première chose que vous allez notez çà maintenant çà va être le cours*

86 *Alors la première idée qu'on va toujours avoir à l'esprit c'est que lorsque l'on cherche une probabilité par exemple pour décider une stratégie ou une autre et bien qu'est-ce que c'est la probabilité on voudrait comprendre la probabilité comme une tendance à long terme c'est la tendance à long terme c'est la probabilité elle va nous indiquer cette tendance maintenant pour déterminer la probabilité on peut faire ce qu'on a fait qu'on a approché mais on va définir on va essayer de définir plus précisément un certain nombre de choses la première chose dont on va parler c'est que si on parle de probabilité il faut qu'ait du hasard et la première chose dont on va parler ce qu'on appelle une expérience aléatoire alors une expérience aléatoire c'est un jeu*

87 *C'est à la sortie d'une chaîne de production prendre au hasard les pièces et voir si elles sont conformes au cahier de charges bon qu'est-ce que c'est une expérience aléatoire c'est une expérience dont on ne peut pas prévoir à l'avance le résultat de façon certaine notre jeu on était tout à fait dans ce cas là est-ce qu'on peut prévoir si on obtient pile pile ou si on obtient pile face est-ce qu'on peut prévoir de gagner de dix points ou on si prends deux pièces de gagner vingt points on ne peut pas prévoir on ne peut pas prévoir à l'avance le résultat de façon certaine mais si on ne peut pas prévoir le résultat de façon certaine on peut avoir une idée de tendances de tendances si on effectue le jeu un certain nombre de fois un grand nombre de fois*

88 *Alors il y a quand même une position que je tiens à vous donner nous ce qui va nous intéresser les expériences que nous intéressent c'est des expériences aléatoires on s'intéresse aux expériences dont on peut indiquer l'ensemble de résultats possibles autrement dit on est dans l'incertain parce que on sait pas on ne sait pas le résultat qui peut sortir mais on a une idée de tous les résultats possibles qui peuvent sortir on n'est pas dans l'inconnu sur les éventualités donc on s'intéresse nous aux expériences aléatoires dont on peut indiquer*

Plus gros çà peut être monsieur s'il vous plaît

Çà va être difficile mais je vais essayer je ne garanti rien pas plus que çà

90 *Bon ayant défini ce que c'est une expérience aléatoire dans le cadre de notre travail d'une expérience aléatoire dont on connaît quand même les résultats possibles l'ensemble*

de résultats possibles cet ensemble de résultats possibles on a l'habitude de l'appeler l'univers donc univers qu'on appelle oméga l'univers c'est l'ensemble oméga

91 *De tous les résultats possibles par exemple c'était quoi notre univers c'était tous les résultats possibles qu'on peut obtenir si on lance deux pièces première pièce pile deuxième pièce face deuxième pièce pile pardon première pièce pile deuxième pièce pile on peut obtenir face face aussi on peut obtenir face pile donc nos résultats possibles ici dans notre exemple çà serait oméga pile pile pile face face pile face face alors dans cet univers*

92 *On va définir ce qu'on appelle les événements par exemple je lance deux pièces vous avez lancé deux pièces vous vous intéressez à l'événement j'obtient pile pile parce que vous savez que si vous obtenez pile pile vous obtenez dix points donc cet événement un événement c'est simplement une partie de l'univers et quand je définis l'événement quand je m'intéresse à l'événement obtenir obtenir pile pile obtenir vingt points et bien et bien je définis cet événement c'est une partie de l'univers c'est pile pile alors parmi parmi tous ces événements parmi tous les événements il y a ce qu'on appelle*

Pouvez vous donner un exemple un exemple

Oui si tu veux

93 *On peut donner comme exemple je vais l'appeler quand A je vais l'appeler grand A l'événement qui nous intéressé à nous c'était pile pile çà c'est un événement et c'est bien une partie de l'univers parmi donc ce que nous intéresse de quoi on a besoin de définir si on a besoin de parler d'événement on va parler des événements élémentaires qu'est-ce que c'est un événement élémentaire c'est un événement qui possède un seul élément par exemple donc ici le A est égal à pile pile et bien cet événement d'obtenir pile pile c'est un événement qui est un événement élémentaire alors en adoptant cette manière de raisonner oui en adoptant cette manière de raisonner on va pouvoir çà va nous permettre d'appréhender*

95 *Les jeux d'hasard et en particulier çà va nous aider à décider sur notre jeu sur la stratégie gagnante bon la je ne vous apprend rien de nouveau j'essaie de clarifier des termes des mots que vous avez déjà employés quand vous étiez en première et terminale mais on a besoin pour bien comprendre la suite alors on continue maintenant çà va trop vite*

Bon donc aillant défini çà on va parler de probabilité alors définir une loi de probabilité on va dire la chose suivante on va dire que définir une probabilité définir une probabilité c'est associer à chaque élément d'oméga un nombre réel positif

96 *définir une probabilité c'est çà alors quelle est la particularité de ce nombre c'est que la somme de toutes les probabilités va faire un ici quand on s'intéresse à notre jeu de deux pièces il faut que la somme des probabilités qui sont associées à chaque élément de notre univers et bon la somme fasse un bon la probabilité d'un événement on va la noter p de a çà sera la notation c'est pas le plus important et maintenant pour pouvoir faire des calculs on fera des calculs on sera dans une situation bien particulière dans*

88 *une situations où tous bah les probabilités de chaque événement de chaque événement élémentaire ces probabilités soient toutes égales on parlera d'équiprobabilité et dans ce cadre d'équiprobabilité et bien on pourra faire des calculs sans pour connaître la tendance à long terme pour connaître la tendance à long terme et bien on n'aura pas besoin forcement besoin d'effectuer un grand nombre de tirages mais on pourra faire des calculs a priori alors qu'est-ce que c'est une équiprobabilité l'équiprobabilité çà correspond au cas où tous les événements élémentaires ont la même probabilité et dans ce cas là et bien chaque chaque événement il a comme probabilité la probabilité d'un événement élémentaire c'est un sur le nombre d'éléments d'oméga voilà dans le cas d'équiprobabilité comment on calcul la probabilité dans notre cas de lancers de pièces*

on était dans un cas d'équiprobabilité on a autant de chances d'obtenir pile pile que pile face que face pile que face face et chacun de ces événements donc la probabilité de chacun de ces événements c'est un sur quatre

98 *Et maintenant comment fait on pour trouver la probabilité d'un événement la probabilité d'un certain événement A et bien vous connaissez cette formule pour calculer la probabilité vous faites on vous a appris le nombre de cas possibles divisé par nombre de cas favorables pardon sur le nombre de cas possibles donc en fait çà revient à dire combien d'éléments dans notre événement A et on divise donc le nombre d'éléments d'A par le nombre d'éléments d'oméga pour obtenir quelle est la probabilité d'obtenir pile face bah quelle est la probabilité d'obtenir quand on lance deux pièces de monnaie quelle est la probabilité*

99 *Quelle est la probabilité d'obtenir deux faces différentes*

Un sur tous les éléments

Il y en combien des cas où on a des faces des résultats différents il y a deux cas pile face et face pile et donc la probabilité est d'un demi donc voilà comment depuis d'ailleurs le lycée c'est comme çà que vous savez que vous faites pour calculer des probabilités maintenant ce que je dirait de ce qui paraît important par rapport au temps qu'on a passé à faire le jeu on n'a pas fini on va y revenir parce qu'on s'a toujours pas prononcé sur notre jeu à savoir quelle est la stratégie gagnante

100 *Mais ce qui nous intéresse nous ce qu'il faut bien avoir compris c'est que quand on calcule une probabilité si on n'effectue pas des lancers ce qu'on mesure ce qu'on mesure quand on fait un calcul de probabilité ce qu'on mesure à travers de la probabilité et bien c'est la tendance d'apparition à long terme d'un événement la tendance d'apparition à long terme d'un événement et on a cette possibilité et bien d'effectuer un grand nombre de fois notre expérience aléatoire pour comprendre pour essayer de comprendre cette tendance alors vous avez-vous qu'il nous a fallu pas mal pas mal de temps on a fait de l'ordre on a fait deux cent vingt tirages vous pouvez constatez qu'avec deux cent vingt tirages le nombre qu'on a obtenu la fréquence d'apparition elle dépasse zéro vingt cinq elle était assez éloignée de zéro vingt cinq elle était de zéro vingt neuf mais ce qu'il faut comprendre ce qu'il faut comprendre ce que plus la probabilité c'est vers quoi plus on va faire de tirages donc on parle de tirages à long terme c'est vers quoi à long terme on se rapproche lorsqu'on effectue des tirages un grand nombre de tirages c'est çà qu'il faut avoir c'est bon*

Non

En fait on a écrit mais on n'a pas écouté ce que vous avez dit

Ouais ouais

Je me disais quel calme quel calme

103 *C'est bon vous avez fini d'écrire donc ce que je disais là quand vous étiez en train d'écrire et bien il y a une idée centrale que j'aimeriez que vous aviez en tête à travers le travail qu'on a fait et qu'on va continuer parce qu'il faut se prononcer*

104 *Et on s'est toujours pas prononcé c'est cette idée que quand on fait un calcul de probabilité que vous savez faire parce que le zéro virgule vingt cinq la probabilité zéro virgule vingt cinq il y en a certains qui l'ont trouvé qui l'on écrite dès les première minutes sur les brouillons des calculs que très vite vous avez fait vous avez appliqué ce que vous avez appris au lycée que quand on lance deux pièces on a une probabilité de zéro vingt cinq d'obtenir pile pile mais cette probabilité il faut que vous alliez en tête c'est vers quoi on se rapproche lorsqu'on effectue un grand nombre de tirages c'est çà l'idée c'est la tendance c'est la tendance d'apparition à long terme*

105 *On revient alors à notre lancer ayant maintenant la probabilité quand vous aviez une pièce quelle est la probabilité d'obtenir de gagner d'avoir pile vous m'avez dit c'était un*

demi vous saviez quand vous aviez pile vous pouviez gagner à ce moment-là vous gagnez dix points donc ayant cette probabilité on a fait autre chose on a calculé je gain moyen qu'on pouvait obtenir le gain moyen qu'on pouvait obtenir il est de combien il était de cinq quand on fait quand on joue avec une pièce on peut dire que le gain moyen c'est cinq çà veut pas dire qu'on va gagner cinq on va pas gagner cinq de toute façon on peut gagner dix ou zéro mais si on effectué le jeu un grand nombre de fois et bien notre moyenne de gain se rapprocherait de cinq

106 *Si on joue avec deux pièces vous avez dit que la probabilité d'obtenir pile pile c'est zéro virgule vingt cinq comme lorsque l'on joue avec deux pièces on peut gagner vingt points quel est le gain moyen qu'on va obtenir pour trouver le gain moyen qu'est-ce qu'il suffit de faire zéro virgule vingt cinq on va multiplier ce gain par la probabilité on va faire zéro virgule vingt cinq fois vingt et on va encore trouver cinq çà veut dire que jouer avec une pièce jouer avec deux pièces çà revient çà revient au même par rapport à notre problème qui était de savoir quelle est la stratégie qui va être la plus rapide çà aurait tendance à aller dans le sens de que vous fessiez que jouer avec une deux trois ou quatre pièces çà revient au même quand on joue avec une ou deux pièces c'est pareil çà revient au même bon pour trois et quatre maintenant*

107 *Notre problème est bien analysez ce qui va se passer quand on joue avec trois ou quatre pièces je vous laisse le temps de chercher donc quel raisonnement va-t-on pouvoir faire pour déterminer quel gain quel est le gain moyen lorsqu'on joue avec trois pièces*

Il faut faire pile pile pile pile pile face face face pile comme çà

Faites çà vite fais un raisonnement de type analogue à celui qu'on a fait pour deux pièces alors

108 *Trois pièces quel est le gain moyen quel est le gain moyen*

Une chance sur huit

Oui mais le gain moyen une chance sur huit ou une chance sur neuf

Je ne sais pas

109 *Bon pour les trois pièces quel est le raisonnement qui me le dit oralement Jennys il a trop parlé au début Jennys Sylvain s'il vous plaît on s'écoute comment tu fais avec trois pièces Sylvain*

Bah j'ai réfléchi

C'est bien c'est une bonne façon de prendre le problème comment est-ce qu'on peut faire pour déterminer une probabilité

Les cas possibles

110 *Tous les cas possibles donc si tu viens à notre on cherche les cas possibles et les cas favorables nous il y a combien de cas favorables pour gagner il n'y a qu'un c'est pile pile pile maintenant il y a combien de cas possibles*

on a huit

Il y en a huit ou il y en a neuf

Dix

Six

Tu as fait la liste pour trouver ah tu as pas fait la liste tu as pas fait comme çà

La liste de quoi

Bah la liste de tous les cas

Moi oui

Tu as fait la liste tu l'as fait comment ta liste

Bah j'ai fait pile pile pile pile face pile

Diagramme de Carnot

Un arbre

Il y en a qui pense au diagramme de Carnot on peut faire un arbre on peut faire un arbre *de probabilité vous me dites que votre première pièce*

111 *Donc votre première pièce çà peut être pile çà peut être face votre deuxième pièce çà peut être pile çà peut être face et votre troisième pièce çà peut être pile çà peut être face en tout il peut y avoir combien de branches de l'arbre huit il vient du fait que là on a deux cas et pour chacun des deux cas il y en a deux cas et là il y a encore deux cas en tout il y a combien de cas possibles deux puissance trois et donc la probabilité ici çà va être un sur huit maintenant on va donc là on n'a pas répondu à notre question notre question c'est toujours de savoir avec quel avec combien de pièces on va gagner le plus rapidement possible qu'est qu'on a besoin de faire maintenant qu'on sait*

112 *Qu'on a pile pile pile pour gagner avec une chance sur huit*

Toujours la même chose

Je ne sais pas comment tu sais que c'est toujours la même chose quel raisonnement tu fais

On multiplie par trente

Tu multiplies par trente tu dis que pile pile pile la probabilité d'avoir pile pile pile c'est une chance sur huit et donc tu fais trente fois un huitième pour trouver pour trouver quoi pour trouver le gain moyen et çà fait quoi trente sur huit

Trois virgule …

Çà fait quinze quarts et çà fait combien

trois trois …

Trois virgule soixante quinze

114 *Bon avec quatre pièces donc qu'est-ce que c'est ce trois virgule soixante quinze interprétez-le moi ce trois virgule soixante quinze qu'est qui signifie Jenny signifie quoi le trois virgule soixante quinze*

La probabilité

Non

La moyenne

C'est la moyenne de quoi

De gain

C'est la moyenne de gain alors tu peux gagner trente tu peux gagner zéro mais si tu fais la moyenne sur un grand nombre de lancers si tu regardes la tendance de tes gains alors la moyenne çà va être de l'ordre la moyenne de gains par lancer çà va être trois virgule soixante quinze pour un grand nombre de tirages et çà va être la tendance qu'on va se rapprocher plus on fait de tirages bon et puis avec quatre pièces on trouve vous m'avez dit

Un sur seize

115 *Un sur seize donc quarante*

Quarante sur quarante fois un sixième

Quarante sur seize et çà fait deux virgule cinq bon alors vous vous prononcez comment alors vous vous prononcez comment alors qu'est-ce qu'il faut prendre

Une pièce

Une pièce

Une pièce

Oui c'est cinq

Attendez attend attend attend je n'entend rien là

Une pièce parce que c'est cinq le gain

Quand il y a deux pièces c'est combien le gain

Eh cinq

C'est cinq aussi

une ou deux pièces

Eh oui on peut prendre une ou deux pièces par contre si on en prend quatre on va moins gagner je vous donne un travail je vous laisserais quatre jours pour le faire
Quatre jours
Vous rendez les pièces çà sera un travail à rendre
116 *je vous donnerai les détailles jeudi*

Transcription séance Les circuits

Date : 27 Novembre 2006
N° élèves : 20

Observations :

- Premier colonne : temps [min]
- Enseignant : texte en *italique*
- Elèves : texte en **gras**
- Message adressée à toute la classe : colonne unique
- Message adressée à un ou plusieurs membres de la classe : plusieurs colonnes

2 *Vous allez faire un travail en groupe alors dans ce travail de groupe il va y avoir plusieurs étapes on va travailler sur des on va travailler vous allez travailler sur des circuits électriques on va travailler ensemble sur des circuits électriques qui ont un certain nombre de composants la première chose qu'on va avoir besoin de faire*

3 *C'est la chose suivante je vais vous expliquer vous avez dans chaque dans chaque enveloppe vous avez différents types de composants alors il y aura des composants qui sont des interrupteurs d'accord il y a des composants qui seront des fusibles donc là vous voyez il y a deux boîtes des enveloppes et des fusibles et puis vous avez des composants qui sont des connecteurs des simples des simples fils alors on va on va travailler à partir de ces composants et la première chose que chaque groupe va à voir à faire assez rapidement c'est je vais vous donner de quoi tester chacun des composants qui sont dans chacune des enveloppes alors pour les tester*

4 *Une lampe une ampoule et puis une pile d'accord donc je donne çà à chaque groupe*

5 *Donc à chaque groupe je lui a donné une première enveloppe de composants à tester et ce que je vais vous demande de marquer*

6 *Sur vos enveloppes c'est le nombre le nombre de composants qui marchent qui et le nombre de composants qui marchent pas alors que je me trompe pas*

	Il y a que des connecteurs	**Ils ont mis du scotch dessus**	
	Çà marche		
	De quel côté celui-là	**S'il y a un qui ne marche pas on pourrait le voir**	
7	**Çà marche çà marche**	**Mais non pas comme çà**	
	Revient après sur celui-là	**On a çà pour le tester**	**Et on fait comment çà**
	Celui-là marche pas	**Mais il faut que tu le testes**	*Bah tu testes ceux qui marchent ceux qui ne marchent pas*
	Il n'y a que ne fonctionne pas d'accord	**Çà marche tu as mis un fusible ou pas**	*Et tu marques après sur l'enveloppe combien est-ce qui*

				marchent et combien est-ce qui ne marchent pas
	On a sept oui			**D'accord**
	Oui sept oui		**Trois oui zéro non je répètes**	**Çà marche**
	Et on le remet dans		**Celui-là il ne marche pas**	**Çà marche pas**
	Oui oui tu le remets dans l'enveloppe		**Six oui et trois non**	**Il n'y a un qui marche regarde**
8	*Et vous pour les interrupteurs il faut tester avec marche et arrêt attention*		**En fait çà fait combien en probabilité c'est trois sur six**	
	Marche pas celui-là ne marche pas non plus		**Un demi**	**Çà marche**
	Aucun ne marche il y a quelque chose d'anormal		**Çà fait un tiers**	
	Celui-là il marche il marche		**Oui çà fait un tiers**	
	Tu es sûr		**Oui un tiers mais attends il y avait combien de possibilités il y en avait neuf**	
	Si si		**Tu en avais six que fonctionnaient tu en avais trois que ne fonctionnaient pas donc çà fait un tiers un tiers ne marche pas et deux tiers qui marchent**	*Puis vous me rendez les enveloppes*
9	**Mais attend au lieu de tester tu fais comme çà**		**On met sur le papier qu'il y en a trois qui ne marche pas**	
	Qu'est-ce que tu racontes			
	Oui et si çà clignote çà marche			
10				
11				
12				
13	*On continue maintenant qu'on a testé pour chacune des enveloppes là on sait par*			

exemple que l'enveloppe six l'enveloppe six c'était des interrupteurs il y a avait deux il y avait deux oui et six non dans l'enveloppe sept un oui sept non

14 *Dans l'enveloppe bon etc. etc. Bon maintenant ce que je vous demande de faire je vais avoir besoin d'aide donc je vais garder un testeur j'ai un testeur ici et puis je vais faire je vais réaliser quelques montages alors premier circuit que je vais faire premier circuit que je vais faire je vais utiliser l'enveloppe un et l'enveloppe quatre l'enveloppe un et l'enveloppe quatre tiens tu viens m'aider tu vas prendre un connecteur*

15 *Ah tu prends un connecteur un fusible et puis tu le montes là donc je vais mélanger un peu je mélange je tire un connecteur je mélange je tire un fusible et puis l'assemblage*

Là je vois des trucs très compliqués tu es pas d'accord

Là il y a six oui et tu dis que le premier s'allume

Bon on a fait le premier montage là tu va rester là alors deuxième montage on va prendre d'ici donc je prends l'enveloppe deux et je prends l'enveloppe six

17 *Tiens maintenant c'est toi qui tire au sort bon voilà et donc on va réaliser on va réaliser ce montage le montage suivant attends j'ai besoin de quelque chose attends j'ai besoin de donc*

18

19 *Çà c'est le connecteur que tu as pris dans l'enveloppe*

20 *Donc pour réaliser là ici le circuit en parallèle il a simplement fait le montage suivant on l'a mis comme çà et après il suffirait ici de prendre deux autres connecteurs pour pouvoir le tester donc ici c'est le montage on a un montage en parallèle et puis dernier montage qu'on va réaliser c'est celui-là tu vas prendre l'enveloppe attends que je me trompes pas le six donc troisième montage que tu va réaliser c'est avec ces enveloppes donc l'enveloppe trois*

21 *L'enveloppe cinq et tiens je te ferme les enveloppes alors ce que je vais faire je vais noter sur le transparent pour chacun des circuits là pour chacun des composants je vais notez ce que vous avez trouvé alors donc pour le premier circuit donc enveloppe quatre oui six non trois pour l'enveloppe un oui six et non un*

22 *Donc deuxième circuit dans l'enveloppe deux oui il y en avait deux non il y en avait cinq et dans l'enveloppe six oui il y en avait deux et non il y en avait six et en fin dans les dernières enveloppes alors enveloppe trois oui vous avez trouvé sept non un dans l'enveloppe cinq huit et un et dans l'enveloppe sept sept et un*

23 *Bon alors j'ai entendu j'ai entendu des questions j'ai entendu pourquoi est-ce qu'on fait çà en maths et voilà la question la question qu'on va se poser c'est la suivante je voudrais allumer l'ampoule du testeur avec un de ces trois circuits avec un de ces trois circuits peu importe le voltage peu importe l'intensité du courant je vous demande quel est de ces trois circuits quel est celui que vous me conseilleriez d'utiliser*

Il faut calculer la probabilité que çà marche

24 *Et puis pourquoi pourquoi vous pensez à tel ou tel circuit ce que je vous demande je vous a mis à chacun une feuille transparente de manière à que vous puissiez rédiger répondre à ces question sur papier à la fin bien sûr en groupe en expliquant les raisons de vos choix et puis une fois que le travail soit fait chaque groupe viendra viendra à présenter les idées de chacun d'accord allez est-ce qu'il y a des questions*

C'est un sept qui est marqué pour la boîte trois et sept

Oui sept et un

25 **Je ne sais pas** **Il faut d'abord** **Bon je n'ai pas**

	quoi faire		**écrire sur une feuille**	**d'idées**
	En tout cas ce n'est pas celui du milieu			**Moi non plus je n'en sais rien**
	Je vous a donné une feuille de pour faire votre compte rendu prenez chacun une feuille bien sûr une feuille de recherche une feuille brouillon et ce que je vous conseille comme tout travail de recherche euh s'il te plaît comme tout travail de recherche vous commencez seuls c'est pas en vous regardant dans les yeux que vous allez trouver des idées et vous prenez le temps qu'il faut environ deux quatre minutes pour chaque et commencez librement à comprendre le problème à avoir des idées et puis en suite vous mettez en commun et puis après			
26	**Celui-là est chien**		**Pour moi il faut voir les cas possibles et les cas favorables**	*Commencez seuls votre recherche si vous avez des choses que vous ne comprenez pas demandez aux autres mais il y a forcement besoin de réfléchir seuls d'abord*
27	**B pour moi c'est le B c'est la réponse B** **Ah non je dirais C la réponse est la C**		**Pour moi il faut après voir les chances** *Je vous demande vous noms dessus pour que on puisse regarder après*	*Et après vous commencez à mettre ne commun* **C'est comme tu vois un sur six deux sur cinq comme si a bien marché**
	Mais non C a plus de composants **Ouais** **Non la C a plus de composants** **Oui mais regarde comment ils font** *Le mieux c'est quand même que vous essayez de vous appropriez chacun du problème et puis après vous mettez en commun si il y a des trucs que vous ne comprenez pas du tout vous demandez aux autres mais commencez pour chercher chacun*			

	Moi je penses c'est le trois cinq sept **Moi aussi trois cinq sept**		**Sept un c'est çà que j'ai trouvé**	**C'est le B**
28	**Pour la résistance c'est quelle chance un sur huit**	**Pour moi il vaut mieux prendre la dernière**	**Moi je vais faire çà**	**C'est le B**
	Mais le fusible peut crever entre temps	**Non moi je dirais l'avant dernière**	**Moi pareil**	**C'est le B parce que il y a plus de chances**
	Mais tout dépend du temps de connexion en fait je penses	**Pourquoi**		**Bah non comment tu fais**
29	**Bah ouais**	**Bah tu as huit chances sur cinq que çà marche**		
	Je penses	**Ouais**		**C'est le C qui marche**
	Je penses que si le voltage est trop fort le fusible il crave	**Huit chances sur cinq que çà marche**		**C'est quoi çà**
30		*Dès que vous essayez de prendre une décision c'est qui est important c'est que à partir de dire pourquoi est-ce que vous pensez çà le plus important est d'expliquer les raisons de vos choix*		**Tu mets ceux qui marchent et ceux qui ne marchent pas**
31		**Moi je ne sais pas du tout** **Une chance sur sept une chance sur huit une chance c'est déjà pas mal**		**Pourquoi tu as mis quinze au lieu de Huit et sept**
		Ecoutes il faut faire un arbre çà et çà et çà **Oui un arbre**		**C'est la chance pour que cela marche**
33	**En fait c'est trente huit sur soixante trois**			*Essayez de mettre en commun les idées que vous avez eu commencez à confronter vos*

				différentes idées
	Mais là tu as une chance sur quatre non non pas celui-là	*Vous avez commencé à travailler les différentes idées que vous arrivé à avoir*		**En deux minutes**
	Essayez de mettre tout en commun là bas les différentes idées je crois que le mieux c'est de voir déjà les idées si vous avez déjà progressé	**Il paraît qu'il faut faire un arbre comme le binaire quoi**		**Et qu'est-ce que tu calcules là c'est le nombre de chances**
		Moi je ferais un tableau		**Ouais je compte les oui les non c'est le C**
34	**Pour le deuxième cas au minimum tu as deux chances sur huit minimum** **Attends attends j'explique le deuxième cas au minimum tu as deux chances sur huit bah c'est ce que j'ai compris bah c'est en parallèle et tu as les favorables c'est deux sur huit deux chances sur huit parce que tu as deux bons sur huit mais attends c'est deux chances sur huit que çà marche ou que çà marche pas que çà marche au minimum parce que moi j'ai cinq sur sept de chances j'ai cinq chances sur sept que çà marche pas oui c'est çà bah non regarde**	**Fais un arbre**	**J'ai vu qu'il y a quatre possibilités et j'ai fait les possibilités pour chaque possibilité**	**Le C par rapport au pourcentage il y a vingt deux oui trois non** **Tu calcules le pourcentage**

	et pour le premier tu as quoi tu as fait quoi			
35	**En fait c'est plutôt le dernier**	**Pourquoi vous voulez faire un arbre**	**Comment çà**	**Çà c'est pas bon**
	Le dernier qui marche	**Parce que c'est plus jolie**	**Je part sur la base qu'on a deux composants sur les deux composants on a quatre possibilités et pour chaque possibilité j'ai analysé la proportion qu'on puisse tomber sur la proportionnalité qu'on tombe sur celui qui marche**	*Allez écrivez un peu les idées que vous avez eu en commun*
	Deux cent quatre			**On essaie de déterminer quel est le quel est le**
	Mais moi je ne sais pas si on fait pareil je ne sais pas je ne sais pas comment on peut faire			*Essayez tous ensembles Je sais que c'est pas facile d'expliquer un calcul quand on est dedans*
	Moi j'ai additionné		**J'ai pas d'idée alors pas d'idée**	**Je pense que c'est le C**
	Oui tu additionnes			**Le C**
	C'est pas comme si on aurait un connecteur parce que regarde le premier le deuxième et le troisième çà peut être çà çà peut être çà çà çà çà zéro zéro zéro			*Pourquoi comment tu as fait quel est le raisonnement tu as expliqué aux autres comment tu avais fait*
	Et là tu as quatre sur sept Le premier c'est quatre sur sept			**Non**
36	**Je suis sur le troisième çà ne marche pas ne marche pas**		**Sept un**	*Bah vous expliquez chacun et essayez de dire quelque chose qui représente à tous*

Cinq sur sept je ne vois pas pourquoi		**Je sais que c'est le C mais je ne sais pas pourquoi**	*Chacun présente aux autres sa stratégie comment vous pensez le faire quel raisonnement vous faites quel calcul vous faites pour essayer de trouver une réponse vas-y expliques nous ton truc*
Même moi je ne sais pas pourquoi j'ai fait çà ah si si si si		**Ah tu dis que c'est le C mais comment tu sais que c'est le C tu es quoi vident**	**En fait j'ai pris le nombre de oui te de non par rapport au total çà fait vingt deux oui et trois non et après j'ai fait le pourcentage**
Pourquoi			*Je t'écoutes eh je vais pas dire si c'est bon ou si c'est pas bon çà n'a aucun intérêt c'est plutôt à eux de voir et te dire s'ils ont compris ta méthode et s'ils sont d'accord vous avez compris ce qu'il a fait*
Parce qu'il y a sept fils et il y en a un qui ne marche pas		**Pour moi c'est celui qu'a le plus de probabilité**	**Ouais**
		Mais attends le C c'est et en série et il y a trois composants et dans le B sont deux composants et en série	*Alors qu'est-ce que vous en pensez*
		Non le B est en parallèle	**Il faut calculer le pourcentage de oui**
		Oui en parallèle en parallèle en parallèle la probabilité en parallèle c'est une chance sur deux parce que si tu as un que ne marche	**Le pourcentage sur le C quatre vingt pourcent de oui**

			pas tu as l'autre qui peut marcher	
	Et moi j'ai raisonné à l'inverse		**Et là tu as six qui marchent et un qui ne marche pas en parallèle**	**Le C Le pourcentage de oui**
	Mais c'est pareil		**Imaginons que tu as un qui marche ici et qui marche pas ici**	**Non moi j'ai fait j'ai fait les oui par rapport au total**
	Je ne sais pas trop en tout cas je sais que c'est la dernière celle-là a un quart de chances		**Là il faut que le deux marchent**	**Et toi tu as quatre vingt oui sur et après j'ai mis çà ne pourcentage mais**
37	**Là tu as deux chances sur sept**		**En parallèle il suffit qu'un marche fonctionne en parallèle il suffit q'un soit bien quoi fonctionne**	**Ah moi j'ai tout ajouté tous les pourcentages et j'ai divisé par**
	Moi j'ai cinq chances sur sept que çà marche pas bah non que çà marche pas		**Mais sur l'autre il faux les deux**	**Moi j'ai quatre vingt huit aussi**
	Bah oui		**Je ne sais pas**	**C'est à peu près pareil j'ai ajouté les oui sur huit donc j'ai mis çà sous un même dénominateur**
	Tu as moins de cinq		**Je ne sais pas mais là sur sept il n'y a qu'un qui ne marche pas**	**Je ne sais pas**
	Bah non		**Et là c'est pareil sur celui-là imaginons qu'il y a un qui ne marche pas et il est foutu**	**Bah moi je ne sais rien non plus mais il est plus fort**
	Bah si		**Je ne sais pas**	**Donc on a quatre vingt huit**
	Tu as l'inverse		**Bon alors on fait quoi**	**Moi j'ai eu quatre vingt sept**
	Bah non que çà marche pas tu as cinq chances sur huit		**Je ne sais pas j'ai envie de dire que c'est le B qui fonctionne bien**	**Tu es sûr**

38	**Mais de toute façon c'est ici c'est le plus défavorable**		**Moi aussi**	**Il y a un problème là**
	Tu as deux chances sur sept que çà marche		**Mais voilà mais peut être le C aussi il faut voir**	**Pour moi c'est celui-là**
	Et là çà je ne sais pas comment tu as fait si tu les additionnes çà fait		**Tu as une possibilité quand même**	**Pour celui-là tu as vingt cent pour cent par là et vingt huit pour cent de chances vingt cent et quarante cent pour cent par là et tu as vingt huit pour cent par là**
	Çà fait vingt cinq composants vingt cinq composants tu as vingt cinq composants tu as vingt cinq composants et tu en as quatorze		**Bah on note un truc ou pas**	**parce que tu es pas obligé d'avoir les deux**
	Cinquante six		**Notes je ne sais pas**	**Lesquels**
39	**Il y en a que vingt deux qui fonctionnent**		**Personne ne décide**	**Les deux**
	Tu dois avoir trois qui ne fonctionnent pas		**Cinq plus sept plus sept sur huit**	**Ouais mais de toute façon c'est le trois**
	Trois sur vingt cinq		**Non non non je réfléchis çà çà marche pas**	**Non mais là tu as deux chances que passe par-dessus tu as deux sur six que çà passe par-dessus**
	Oui trois composants sur vingt cinq		**Pour moi il faut les multiplier les multiplier quand c'est en série et le diviser quand c'est en parallèle**	**Parce que tu vas à avoir**
	Non c'est pas çà		**Ouais bon bas écoutes on va regarder**	**Bah c'est pas deux sur six**
40	**Tu as trois composantes qui ne marchent pas c'est pas comme çà qu'il faut faire**		**On va voir qu'est que çà va donner**	**Non non deux sur quatre qui marche**

	C'est compliqué le trois		**Quoique si tu additionnes bah non si tu additionnes çà marche pas**	**Là tu as fais deux sur six**
41	**Tu as combien**		**Si tu additionnes les cas où çà marche**	**C'est pas deux sur six deux sur huit un sur quatre que marche**
	Quatre sur sept et quatre vingt cinq sur sept et le deux tu as quatre vingt cinq sur sept		**Pour moi c'est toujours B**	**deux sur quatre qui marche deux sur quatre**
	Non non c'est pas comme çà		**Parce que la C non seulement ils sont serrés il y en a trois dans la C en plus**	**Tu as fait deux sur six mais c'est deux sur huit qui marche il y a deux oui et six non c'est pas deux sur six et tu fais le calcul çà fait vingt cent pour cent là c'est oui plus non pour avoir le total**
	Moi j'ai fait j'ai fait tous		**Regarde sur vingt deux composants tu en as que trois qui sont défaillants**	**Le truc c'est que là là j'en ai trois**
	Ah tu as fait toute la liste de		**Il y en a que trois qui sont défaillants**	**Le truc c'est que là tu as trois**
	Ouais		**Bon on a fait çà et après**	**Là je suis d'accord je suis d'accord tu as la meilleur probabilité mais si tu tombes sur un là tu tombes sur un que rien ne marche donc**
	C'est pas pareil		**On pense que c'est le B**	*Vous pouvez me diter un peu où vous êtes*
	Je suis certain		*Et pourquoi*	**On tourne en rond**
	Là tu as quatre sur sept		**On ne sait pas mais c'est celui-là**	**On a fait tous les pourcentages de oui et de non tout le monde est là mais après on s'est dit que c'est le C qui marche s'il y a un seul élément sur les**

	Çà fait pas quatre sur sept çà fait çà zéro cinquante sept			**trois Qu'il y a quand même une chance sur huit à chaque fois ou sur neuf que qui tombe que le circuit qu'alors que le B par exemple il y a deux chances sur sept et comme ils sont en dérivation on a un que çà peut pas marcher et il marchera quand même**
42	**Tu as écris tout çà**		*Le plus important comme en tout problème de maths c'est pas de savoir bah si c'est important mais c'est aussi important qu'est-ce que vous donnez comme raisons pour quoi vous allez choisir B ou C*	*Oui*
	Ouais	*Vous êtes où dans vos idées*	**En fait on ne sait pas s'il faut additionner ou multiplier ou soustraire ou intégrer**	**Donc voilà**
	Zéro cinquante sept tu as trouvé	**En fait j'ai pas beaucoup développé bah je ne pourrais pas dire pourquoi est-ce que je ne sais pas c'est toujours un sur mâchant un sur mâchant c'est ce qu'il faut après prouver je ne sais**	**On a deux probabilités**	*Donc vous n'arrivez pas à vous prononcer quoi*
	Quatre vingt combinaisons quatre vingt combinaisons pour l'un	*Bon mais quand même essayez de rédiger pour discuter après avec les autres parce que*	**Une loi de poisson**	**Non**

		pas tous ont pensé comme vous par exemple		
	Mais çà te fais zéro cinquante sept		**Une loi de poisson**	*Bon vous mettez quand même vos premières idées que vous avez eu y compris si vous êtes coincés et pourquoi vous êtes coincés*
	Eh ouais		**On verra çà après je ne sais pas qu'est-ce que c'est**	**Pour le B là çà fait quatre oui et onze non**
	Ah c'est bon je peux le faire maintenant		**Qu'est-ce que c'est çà**	**Oui mais çà fait quatre chances sur quinze**
	Comment tu fais pour trouver çà		**Moi je sais que pour deux composants il faut faire oui oui oui non non non non oui**	**Tu peux avoir deux chances sur sept que çà marche**
	Qu'est-ce qu'il y a		**D'accord tu fais six oui et trois non fais un oui et deux non je ne sais pas**	**Deux chances sur**
43	**Donc c'est l'A qui marche bien bah pour moi c'est l'A qui marche bien tu as trouvé un qui marche**		**Mais attends çà veut dire que pour le parallèle çà fait**	**Deux chances sur sept pour le B que çà marche deux chances sur six que çà marche dans le sens qu'il peut y avoir**
	Le C pour moi		**Pour le parallèle çà devrait**	**Le tout çà fait cinq plus deux**
	Qu'est-ce que vous avez mise en commun vous arrêtez vous calculs alors chacun va exposer sa méthode ce qu'il a fait		**Pour le parallèle tu as celui-là celui-là celui-là et pour le C tu as**	**Sept**
	En fait j'ai multiplié le nombre de sur le premier cas par exemple il y a six composants sur sept qui		**Je ne sais pas pour quoi mais je sens qu'on va mal**	**Pour trouver**

fonctionnent et six composants sur neuf j'ai multiplié les deux j'ai eu trente six sur cinquante trois et çà me donne cinquante sept pour cent		
Toi tu as fait pareil	**Tu as plus de chances que celui-là marche**	**Donc on a deux chances sur sept que çà marche**
Ouais	**Là tu as quinze chances sur vingt huit pour tomber le cas qui marche**	**Non le truc c'est que tu as aussi**
Moi j'ai fait plus compliqué moi	**Il faut pas les additionner il faut les multiplier**	**Deux chances sur huit que çà marche**
C'est comme si tu avais multiplié	**Mais combien des possibilités tu as là**	**Et il peut y avoir des truc qui marchent et d'autres qui ne marchent pas**
C'est-à-dire qu'est-ce que tu as fait	**Neuf**	**D'accord c'est çà**
J'ai fait une sorte de binaire mais	**Neuf**	**Çà on est d'accord sur le truc au niveau de pourcentage au niveau de pourcentage c'est plus bas sauf que là là tu en as un un sur**
Ouais	**Bah je vois pas oui oui oui et non non non après oui oui**	**Un sur sept qui peut marcher**
J'ai trouvé un peu çà	**Et là c'est non**	**Oui là tu en as un sur sept qui marche par là et tu as deux sur huit qui marchent par là et le truc c'est que tu es pas indépendant des autres quoi**
Et moi j'ai trouvé pareil	**Il y a forcement autant de non que de oui**	**Là tu en as un sur les trois un sur les trois qui fonctionnent**

		Soyons bref pour celui-là dès qu'il y a un non çà marche pas et regarde tu as pas celui-là tu as pas celui-là tu as pas celui-là	**Là on additionne on additionne tout tu auras un sur qui va faire chier**
Essayez dans votre transparent dites les différentes manières que vous avez utilisées résumez les différentes travails et après si toi tu additionnes et toi tu multiplies peut être toi tu as raison et lui il a tort ou au contraire je ne sais pas il va falloir donner les arguments et puis si il y a des choses que vous ne savez pas vous les mettez aussi dans votre compte rendu			
		Celui-là il ne marche que si il y a les trois	**Mais là tu as plus de risque en pourcentage**
		Donc il y a plus des possibilités avec celui là	**Quelle chance tu as que çà marche et quelle chance tu as que çà déconne**
		Mais attends c'est par rapport aux composants qu'il faut voir la possibilité	**Pour que çà déconne il faut le mettre sous le même dénominateur**
		Ce truc me dépasse	**Non non là on revient tu vois tout de suite que non non çà c'est un cas à part çà là et là c'est un cas à part entre là et là c'est un cas à part**
Je pense qu'il faut additionner parce que j'ai tort par rapport à toi en maths		**Là j'ai fait n'importe quoi**	*Vous avez commencé à faire votre compte rendu*
		C'est huit c'est deux au cube	**Mais on mets quoi**

45	**D'accord tu as tort et il a mis comme moi**			*Bah quelles sont les raisons quelles sont les raisons de votre choix*
	Moi j'ai faut			**On met juste les pourcentages**
	Mais tu as fait comme moi tu as zéro cinquante sept			
	J'ai faux j'ai réfléchi tu as plus de composants			**Si on regarde çà quel est le pourcentage que çà passe par là et quel est le pourcentage que çà passe par là là on peut le savoir mais là sur les trois foutus ceux qui ne marchent pas**
	Mais après il faut le multiplier c'est comme les chances quoi			
	Attends on va faire quelque chose pas compliquée			**Là si tu tombes sur un qui ne marche pas rien ne marche**
	Mais attends çà y est	**Il faut revenir**		
	Mais non moi c'est pour trouver c'est juste çà il y a un sur deux et l'autre un sur deux	**Il faut comprendre pour quoi pour savoir comment réfléchir quoi**		**Ouais**
	Attends où est marqué çà dessus			
46	**Et pour les deux il y a une chance sur quatre**	**Il faut faire le binaire quoi**		**Tu es d'accord tandis que celui-là que ne marche pas te celui-là qui marche**
	Pour le un tu as une chance sur deux et pour le quatre tu as une chance sur deux çà fait une chance sur quatre tu vois			

	là il faut multiplier il faut multiplier les trois			
	On a six composants sur sept	Un tableau je pense		Là j'ai fait j'ai additionné les deux et j'ai divisé
	Tu as combien	Mais tu fais comment le tableau		
	Trois mille quatre vingt seize	Un arbre		Vingt huit et vingt cinq vingt huit d'un côté et vingt cinq de l'autre
	Vas-y			
	Et sur sept composants trois mille quatre vingt seize sur quatre mille cinq cent …			C'est pas bon en fait
47	Que fonctionnent tu as trois mille quatre vingt seize			
	Donc on multiplie			Eh
	Zéro soixante huit			
	Moi j'ai six sur neuf donc on multiplie			C'est pas bonne comme çà
	Toi tu multiples			
	On multiplie			
48	Vas-y mets pour la première quand même			Non mais là tu as vingt cent pour cent de ce côté que çà marche et vingt huit pur cent de l'autre
	C'est zéro cinquante sept			
49	Cinquante sept pour cent	Ecoutes à gauche tu prennes tout ce qui est foutu à droite ceux qui sont bons	Déjà ce n'est pas la C	Ouais
	Moi j'ai quarante sept pour cent	Qu'est-ce que tu racontes	Pour la première possibilité tu as six chances sur vingt et un de tomber sur un composant	
	Eh	Bah oui comme çà tu as les possibilités	Eh arrêtes les conneries comment tu fais pour faire tout çà sans la	Là tu es d'accord que là

		calculette	
Moi j'ai pris le mauvais côté il faut prendre neuf d'un côté et sept de l'autre			
Que moi j'ai quarante sept pour cent	**C'est pas possible arrêtes de déconner**	**A moi çà me suffit comme çà je l'ai fait à la main**	
Mais lequel finalement l'A le B ou le C	**Mais non tu vas voir**	**Expliques moi**	**Là tu peux additionner les deux tu as cinquante trois pour cent des chances que çà marche**
Attends explique pourquoi tu as fait comme çà	**Regarde**	**Je commence à m'énerver déjà**	
Attends que je fasse les autres après je t'explique	**Il y a trois chiffres c'est çà**	**Avec la calculette tu veux quoi**	**Non moins que çà**
Tu as six composants sur sept et après tu as six composants sur neuf	**Non c'est pas çà c'est pas sur vingt deux**	**Non trente six soixante troisième**	
Tu as fait comment çà expliques			
Tu as trouvé combien pour le premier			
Tu as fait comment toi expliques		*Si tu penses que oui tu l'expliques si tu penses que non essayez d'expliquer pourquoi*	**Non**
Et tu trouves comment ton zéro quarante sept	**Il faut diviser c'est çà**	**Bah oui**	**Tu as vingt six pour cent**
Cinquante sept	**Donc il y a une chance sur sept**	*Revient au sens que tu as des probabilités tu as des raisons de faire çà et pourquoi tu penses qu'il faut faire çà*	**Non**
Peu être je me	**Mais c'est quoi çà**		**De chances**

	suis trompé	**une chance sur sept**		
	Attends je dois faire pareil et j'ai fait pas pareil			**Non**
50	**Mais attends c'est un truc en parallèle c'est rien à voir**			
	C'est en parallèle c'est pas comme par rapport à deux septième			**Ah oui oui oui parce que tu peux pas additionner les deux**
	On néglige deux huitième on ne compte deux septième			
51	**C'est en parallèle c'est rien à voir**			**Parce que là**
	Refais plutôt l'un			
	Ce que j'aimerais il vous reste cinq minutes donc vous faites dans les cinq minutes faites votre compte rendu des idées que vous avez eu je vous rappelle la question c'est qu'est-ce que vous me conseillez et puis vous me donnez des raisons pourquoi vous me conseillez çà et vous devez aussi dans votre compte rendu dire les différentes idées d'accord			
52	**Je ne sais pas comment faire la démonstration**	**C'est le deuxième qui marche**	**Et bon je marque quoi là les gars**	**Je suis d'accord qu'au niveau de pourcentage je ne sais pas mais**
	Je peux te montrer je te montre là c'est un demi une chance sur deux	**Non mais c'est le tableau ce que je dis**	**Pour moi c'est entre A et B**	**Attends on va mettre cinquante trois pour cent là**
	Moi je ne comprend rien de ce tu as marqué	*Qui va parler là*	**Alors je marque quoi je marque quoi**	**Mais tu vas l'expliquer**
	Regarde là si tu additionnes c'est deux chances sur quatre	**C'est lui on a voté déjà on n'a pas unanimité**	**C'est le B qui marche c'est entre les deux**	
	Bah oui mais là tu dis n'importe quoi	**Le B est en parallèle**	**Moi je dirais le B c'est le B regarde**	
	Là tu as quand même je ne comprend plus rien là	**Et celui-là de tout façon çà marche toujours**	**Pour moi c'est le A**	
	Attends si tu le laisses ouvert	**Mais pour le deux je ne sais pas comment faire pour savoir si çà marche**	**Théoriquement si tout marchait s'il n'y avait pas des oui et de non si çà**	**Le B tu as mis combien que çà marche**

		ou si çà marche pas	**serait le B avec B on a plus de chances de tomber sur un qui n'est pas bon c'est en parallèle alors**	
53	*Là vous mettez le nom et pour le brouillon c'est pareil même si c'est pas très propre c'est juste pour voir les idées*	**Il est pas mail mon tableau çà peut marcher pour le premier plutôt**	**En fait çà n'a rien à voir que soit en parallèle**	**Tu dis que le C c'est le meilleur montage**
54	**Il faut multiplier et pas additionner**	*Et là vous expliquez maintenant vous expliquez votre décision comment vous avez fait quelles sont vos différentes idées*	**Ecris le A tu dis qu'on va choisir le A moi je dis le A**	
	Pourquoi tu as fait comme çà **Bah pareil que pour la première** **Mais il n'y a aucune certitude aucun des cas** **Mais non aucune** **Si tu additionnes ou si tu multiplies on n'est pas sûr**			
55	**Et monsieur vous pouvez pas nous dire si c'est bon ou pas**	**Pour moi le B**	**Il faut qu'on choisisse**	**Alors pour le montage B comment ils ont fait là eux**
	Après on essayera d'établir d'essayer de comprendre ce qui c'est la probabilité et après je pense que vous finissez votre compte rendu et puis vous prenez un peu de l'air vous restez dans el couloir **Je fais la démonstration là** **Je met combien**	**Pour moi aussi le B**	**Là tu peux faire un petit tableau un schéma et tu peux rentrer les probabilités**	

56	**Notre truc est complètement faux**	**Amis alors si les deux sont morts**	**Pour moi c'est le B ah l'A d'accord**	
	Il y a quatre composants	**Bah se les deux sont morts**	**Là c'est un septième**	**Là on a deux sur sept deux sur sept**
	Là tu as deux composants une chance sur deux	**Bah je n'ai rien compris de rien je n'ai rien trouvé**	**Çà c'est nul**	
	Une chance sur deux et une chance sur deux égal à une chance sur quatre car	**Tu dis tous les composants morts c'est çà**	**Bah si tu fais mieux tu écris**	
	Comment ils ont trouvé les autres groupes			
	Je sais pas			
57	**Je ne comprends pas ce que tu as fait**	**Je ne pourrais pas l'expliquer çà c'est sur**	**Oui parce qu'on commence à en avoir marre**	**Donc là çà fait on a dit vingt huit pour cent**
	Là tu dis je ne comprends pas et tu m'as dit que tu avais compris		**Non non je ne sais pas quoi dire**	
	J'ai compris ton idée principale mais là	**Et Jérôme là il a fait quoi**	**On se calme d'accord**	
	Mais c'est pareil	**Bah il dit qu'il faut voir**	**Tu dis qu'il y a quatre possibilités**	*Çà y est vous avez fait votre compte rendu et là vous dites quoi là*
	Ouais je sais pas	**Bah je ne sais pas**	**Je ne sais plus comment faire**	
			C'est bon j'ai trouvé pour le truc du milieu il y a …chances que çà marche et …que çà marche pas	**On va écrire en pourcentage la meilleure connexion**
			Mais attends il faut dire quel truc choisir c'est le meilleur qu'il faut choisir il faut que tu expliques pourquoi tu as choisi celui-là	
			Cà c'est le boulot	
58	*Bon on fait une petite pause de cinq minutes*			
	Là je mettrais		**Tu vois je ne sais**	

	bien en rouge là il y a deux composants il y a un qui fonctionne l'autre ne fonctionne pas	**pas s'on multiple**	
59	**Tu fais la C là**	**Et Jenny il multiplie ou il additionne**	**Le C a un pourcentage meilleur mais le B**
	Là je trouve soixante seize pour cent	**Je ne sais pas**	
	Que çà marche	**Alors tu fais çà plus çà moins la probabilité de A fois B**	
	Ouais donc çà veut dire	**Ok, après il faut savoir quelle est laquelle**	
	Soixante seize pour cent moi j'ai trouvé soixante huit pour cent	**Huit non çà veut dire que celle-là fonctionne**	**Pour moi c'est le B**
	De toute façon c'est la dernière pour moi c'est la dernière	**Il faut calculer la probabilité pour chaque et après tu multiplies**	**Mais il faut l'expliquer**
60		**Fais le premier et çà fait trois six sur**	
61	*Je vous laisse deux trois minutes pour terminer le compte rendu*		
	Qui présente là		
	Moi j'ai fait çà j'ai écrit	**Mais on les additionne ou on les multiplie**	
	D'accord toi tu as fait le transparent d'autre qui présente		
	Non moi j'ai pas les mêmes idées que lui pas moi	**On le multiplie çà fait trente six sur**	
	Mais il va falloir que tu l'exposes ton idée	**Attends je le refais**	
	Parce que toi tu as multiplié		
	J'ai pas compris		
	Non bah j'ai fait trop de brouillon moi		

	Parce que toi tu as multiplié Et moi j'ai additionné		
62	**Moi je ne peux pas commenter c'est clair**	**Il faut voir celui de plus de probabilité**	
	Mais justement si je présents je suis pas convainquant	**Pour le premier cas il y a combien de chances**	**Six divisé par sept multiplié par six divisé par neuf quatre septièmes**
63	**Donc tu as soixante dix pour cent çà fait combien çà**	**Douze chances que çà marche**	**Ah tu fais un truc pour chacun ah oui oui ah**
	Attends je vais regarder combien çà fait toi tu as combien	**Douze**	**Les oui non tu fais trois septièmes**
	Mais tu as additionné tu fais comment pour additionner çà	**Oui douze**	**Ah oui oui d'accord**
	Comme si comme si	**Et pour le deuxième**	**Çà fait trois sur vingt et un et les deux non non les deux non non çà fait vingt septième fois**
	Tu as fait six divisé par sept	*Çà y est çà le termine le compte rendu*	**Alors laissez tomber c'est la pause**
	Un divisé par sept	**Mais tu est sur qu'il y a un tableau pour le trois c'est dur ah**	**Vingt septièmes fois six neuvièmes**
	Tu as fait la chance que ne marche pas moi j'ai fait l'inverse mais c'est pareil	**Fais un arbre çà marche avec un arbre**	**Aha oui**
	Je ne suis pas sur de çà	**Je croyais que c'était cinq bons deux morts**	**Çà fait**
	Tu as six chances sur sept que çà marche et une chance sur sept que çà marche pas	**Et tu as fait comment et là six**	**Çà fait deux sur vingt et un**

	Et le deuxième quarante six pour cent on est bien d'accord	**Alors ouais ouais ouais ouais ouais**	**On continu a faire des choses bêtes je te dis eh**	
	De toute façon c'est le dernier			
	Mais non attend le deuxième c'est bien quarante six pour cent		**Et là çà fait douze sur soixante quinze**	
64	*Bon même si vous n'avez pas fini c'est bon*			
	Et là tu n'as rien écris en plus	**Donc c'est Dennis qui t'a dis de faire le tableau comme çà**	**Je ne comprends pas pourquoi ils font çà**	
	Là c'est pas logique là tu as quatre chances	**C'est pas le vrai çà**	**C'est terminé déjà**	
	Non moi j'ai fait parce que là aussi çà peut marcher bah en fait c'est les mêmes mais là en fait c'est pas pareil là c'est rien à voir là c'est faux		**On a un petit problème**	*Çà y est mais il faut l'expliquer*
	Si	**C'est vingt et un**	*Allez çà se termine le compte rendu*	**On a fait les pourcentages de oui et de non c'est chien d'avoir pour le B vingt sept bah cinquante trois**
	Mais non c'est faux parce que là tu peux avoir là et celui là sur les deux et celui-ci il peut tomber quand même	**C'est vingt sept pour le un vingt sept sur combien il y a combine de bons un deux trois bons et un deux trois quatre six dix huit**	**Attends tout le monde a écrit**	**Mais non c'est évident à première vue c'est le B**
	Mais çà c'est faux c'est faux c'est sûr		**Il y a un truc qui ne fonctionne pas**	
	Peu concluant je le vois	**Mais quel ordre ABC ou BCA ou**	**Et tu as fais toute l'addition tu fais trois plus six neuf et neuf plus douze**	
65	**Çà c'est faux regarde**	**Trente deux chances que soit**		

	bon pour l'a pour l'a il y a trente deux chances qui soit bon et il y a vingt sept qui soit pas bon
En fin de compte non	**Et pour le b**
Là çà c'est faux j'en suis sur tien regarde là tu peux tu peux tomber sur celui-là aussi	**Bah on n'a pas de temps quoi**
Pour moi l'ordre est CBA	

Je vais vous demander maintenant de m'écouter s'il vous plait

CBA CBA	**Pour le B pour le C je dis on ne peut pas faire le tableau**

Non CAB
CAB pour moi

Tiens on va commencer par Roman là tu nous présentes ce que vous avez fait et ce que tu me conseilles de faire

Donc on a fait le pourcentage de oui et de non pour les trois montages pour le premier on trouve douze oui et quatre non çà fait soixante quinze pour cent de chances que çà marche pour le B on trouve quatre oui et onze vingt sept pour cent que çà marche sur le C vingt deux oui et trois non donc quatre vingt huit pour cent que çà marche le montage C pour moi c'est cinquante trois parce le B que on a vingt huit pour cent que çà passe par là haut et vingt cent pour cent pour en bas don l'addition j'ai additionné tout çà

Je sui d'accord pour le B	**Çà revient au même**	**On va essayer de le faire avec çà maintenant**

Est-ce que vous avez des questions des choses que vous n'avez pas comprises moi je sais par rapport aux démarches que vous avez que j'au vues je sais pas probablement il y a des questions à poser ouais

Pour le cas A et C je sais pas comment vous trouvez un pourcentage plus élevé que les composants ce soixante quinze pour cent….pour les fils et … pour les fusibles

Je pense qu'il y a une chose que tu n'as pas bien compris …c'était oui six non un c'était pas soixante un pour cent ici là c'est bien de poser une question le oui du six c'était six il y a six oui et trois non finalement vous nous conseillez le

Le C

D'accord allez deuxième groupe il y en a qui nous présente qui nous dit comment vous avez fait que nous conseilleriez

Bah en fait pour le premier et le dernier cas on a multiplié et le dernier groupe bah c'est ce qu'on a fait donc on a multiplié les chances et par exemple on a

	Il y a deux qui marchent

Eh attend attend écoutez tous les points de vue sont intéressantes écoutez-les et après on *discutera si on n'est pas d'accord vas-y*

Donc pour le premier cas on a deux composants le premier composants bah la première sorte de composants on a une chance sur deux et l'autre une chance sur deux donc si on multiplie on a une chance sur quatre donc même si le premier composant soit il est ouvert soit il est fermé donc çà çà fait une chance sur quatre donc c'est pour çà que pour le premier cas et pour le troisième cas on a dis de multiplier les chances et çà nous donne cinquante sept pour cent et soixante huit pour cent et après pour le deuxième cas on ne sait pas

Tu vois rien à voir

Il faut l'écrire c'est une chance sur deux

Et c'est quoi ce sept

C'est la proportion

Alors je ne présente pas les gars

Des questions troisième groupe vous voyez comment c'est intéressant de confronter les points de vue sont intéressants

Tu as vu comment ils ont multiplié

C'est bien si on multiplie

Ce sont les cas tu vois qui fonctionnent sur tous les cas

Oui j'ai vu il faut l'écrire

Cinq sur vingt huit quinze sur vingt huit

69 **Nous pour les trois cas on a fait des tableaux en fait on n'a pas fait grande chose on n'a pas trouvée une méthode à la fin on a fait des tableaux pour le B pour le B pour le B donc on a douze cas que çà marche et dix huit que ne marche pas**

C'est trois quarts

Bah c'est bien quinze sur vingt huit quinze sur vingt huit c'est bien

Oui vas-y

Çà me fait peur

70 **Pour le cas A on fait pareil on prend ce qui est possible et ce qui n'est pas possible on trouve vingt sept cas négatifs et trente deux cas positifs mais je ne sais pas comment expliquer çà**

Déjà j'ai trouvé les douze cas que çà marche et les dix huit que çà marche pas

Mais çà marche pas le deuxième

C'est trente cinq

Donc vous avez des questions à lui poser vous avez des questions moi j'aurais une question est-ce que vous pouvez est-ce que vous pourriez l'un d'entre vous avec un marqueur pourriez me montrer plus précisément ce que vous entendez par attend attend il va nous expliquer comment il a fait sont tableau peut être çà va nous aider

Oui moi aussi j'ai une question là-dessus

Le tableau c'est pas le même

Fais le directement là-dessus là

Bah on a comptez les cas là du tableau

Aha d'accord groupe suivant

Il faut faire çà plus propre

Jeremy je te laisse la place

Oui c'est toi qui a eu l'idée

Le plus jeune

72 **Pour le premier cas on a pris la formule a fois b et on a trouvé cinquante sept pour cent de chances qui marche on a multiplié les pourcentages des chances que les éléments fonctionnent et pour le troisième cas c'est exactement pareil mais comme il y a trois composants il faut multiplier les trois pourcentages et on obtient soixante huit pour cent et pour le deuxième cas on a fait la formule de a o b on prend les cas où on a a et les cas où on a b et soustraire les éléments de a et b c'est comme des éléments en plus et on trouve quarante six pour cent des chances quarante six pour cent de chances que fonctionne le montage**

73

74 *Des questions comment si si vas-y*

D'où vient la formule

Où tu as trouvé la formule

C'est la formule qu'on a apprise de a et b et d'a ou b

C'est une formule automatique

Si j'aurais eu cette formule j'aurais fait comme çà

Des réactions là vous en pensez quoi de cette formule là les autres

Pour moi c'es bon

Vous en pensez quoi de cette formule c'est une formule que vous avez apprise

On a appris une formule mais je ne sais pas si c'est celle là

Moi j'ai jamais appris çà

Moi jamais j'aurais utilisé çà

J'ai appris une formule mais je ne sais pas si c'est celle-là

Il faut voir pour que çà marche

Que çà marche pas

Dans notre tableau oui c'est çà marche

Eh

Bah oui

Il y a une chance sur huit que çà marche pas

Bah le oui c'est le O

Là çà marche pas

Bon alors vous passez vous nous montrez
Mais çà sert à rien cette formule
Ces gars ils ont déjà cette formule
Quelle est la probabilité que c'est que tu as fait soit bon
75 **Mais çà sert à rien**
Tu as fais un truc compliqué
Tu va a avoir faut mais ne t'inquiètes pas
J'aurai eu de changer ces chances
Vous écoutez tous
On a fait la probabilité pour chaque …
Parle un peu plus fort moi je t'entend mais les autres pas
Est-ce que vous avez des questions là comment
On n'a rien entendu
Tu n'as rien entendu vas-y tu nous expliques tourne plutôt vers là bas et parle plus fort
…on a pris le cas qui a la plus forte probabilité de fonctionner bah par rapport aux autres en fonction des calculs
Et c'est quoi les calculs c'est çà un peut la question qu'on a
En fait on a pris toutes les possibilités qu'on avait pour chaque possibilité et sur chaque possibilité qu'on avait et on a trouvé la probabilité pour chaque
En les multipliant
Un tableau de carnot
On a pris les possibilités qu'on avait et sur chaque possibilité et on a pris les pièces qui étaient bonnes et les pièces qui étaient pas bonnes et par les composants
Et le vingt trois sur soixante douze c'est quoi
Le vingt trois sur soixante douze c'est où là
En bas
Quarante neuf sur soixante douze c'est oui qui fonctionne et vingt trois sur soixante douze c'est la possibilité que ne fonctionne pas
Et comment vous avez fait pour trouver ce quarante neuf sur soixante douze
78 **Bah on avait trois oui et on a multiplié chaque possibilité on va dire et on l'a appliqué on a multiplié et pour le vingt trois sur soixante douze on a fait un moins vingt trois sur soixante douze**
Tu as trouvé qu'il y a une chance de vingt trois chances sur soixante douze que le circuit marche
Non que ne fonctionne pas
Qui ne marche pas
Ouais
Et tu as trouvé quarante neuf sur soixante douze que çà marche tiens et toi tu prend ta calculatrice pour faire quoi
Bah non non comme çà
C'est pour le mettre en pourcentage
Ouais c'est pour le pourcentage
Si si si si je pense que tu l'as pris pour faire quelque chose

A voir ce calcul là
Pour le voir en pourcentage en pour cent
Quarante neuf sur soixante douze çà fait soixante huit pour cent à peu près
Donc là tu as trouvé soixante huit pour cent

C'est bon on a ce numéro	**Bah comme nous quoi** **C'est la même chose bah c'est moi qui l'a fait**

…bon alors je vous propose la tâche çà va être difficile on va essayer de reconstituer les morceaux du puzzle je vais essayer vous arrêtez maintenant ah vous alliez nous expliquer comment vous avez fait le tableau

80 **Donc voilà pour le cas numéro un il y a un des éléments bah qu'il y a six que sont bons et trois qui ne sont pas bons et deuxième élément il y en a six que sont bons et un qui sont pas bons donc on fait tous les cas avec tous les éléments six bons et les mauvais il y a six bons et trois mauvais et après on fait le tableau et en trouvant tous les cas o qui sont bons et les cas croix qui ne sont pas bons çà fait vingt sept pas bons et trente six bons et on fait le pourcentage**
Ouais ouais
Est-ce que çà contredit les choses qu'on été dites là sur une feuille sur une feuille je vois cinquante sept pour cent là c'est cinquante huit que je trouve peut être qu'il y a un problème de méthode je ne sais pas bon alors je vais qu'on vous voyez j'ai tout démonté j'ai tout démonté j'ai tout mis dans les j'ai tout mis dans les enveloppes et je vais vous poser une question déjà un petit peu plus simple pour qu'on soit bien d'accord sur pour qu'on soit bien d'accord sur comment on prend si cette idée de cette idée de décision à prendre bon alors je vais prendre je commence par je vais prendre la première enveloppe je prend la première enveloppe donc la première enveloppe je remets mon transparent de départ donc la première enveloppe
Ok cinquante sept de bons et quarante deux des pas bons est-ce que çà est-ce que çà contredit les choses qu'on été trouvées non ah je vois sur une feuille je vois sur une feuille pour le premier cas je vois un cinquante sept pour cent là c'est cinquante huit que je trouve c'est un problème de précision

81

82 *Donc première enveloppe où on avait six oui un non je tire au hasard un connecteur et je le mets ici je veux prendre la quatrième enveloppe des fusibles je tire au hasard un fusible je le mets ici et puis je veux faire çà pour chacune des enveloppes alors je vais continuer et en suite c'est quoi c'est l'enveloppe deux*
Ah oui c'est la cinquième j'en étais où la c'était la six bon je vais continuer et puis vous allez me dire je voudrais maintenant j'ai fait çà pour chaque j'ai fait çà pour chaque chacune de mes sept enveloppes je voudrai que mon système que ma lampe que mon ampoule s'allume vous me conseilleriez de prendre lequel parmi ces sept que j'ai sortie
Le cinq
J'en prend une et je voudrais que mon système je voudrais qu'avec mon testeur çà marche
Bah le cinq

Mais on ne peut pas savoir
Bah mais tu fais celui qu'a plus de chances tu as plus

de chances avec la cinq

De la pochette cinq

Attends tu dis quoi

Qu'on ne peut pas savoir comme lequel marche

On ne peut pas savoir pourquoi tu ne peux pas savoir

Bah bien bien sûr

Et toi tu penses quoi toi

Bah sinon celui de plus haute probabilité là on a une chance sur neuf donc la pochette cinq

Donc là tu conseillez la pochette cinq tu me conseillez la pochette cinq qu'est-ce que fait pourquoi tu me conseillez la pochette cinq qu'est-ce qu'il fait pourquoi tu me conseilles la pochette cinq

Bah il y a quand même un nombre important d'oui par rapport au nombre de non

Donc finalement qu'est-ce que tu vas à calculer ici c'est la proportion

Voilà c'est çà

La proportion d'oui par rapport au nombre total de composants qu'il y a dans l'enveloppe

Oui oui c'est çà

D'accord donc maintenant ce que nous dit Dennis ouais mais je prends de la pochette cinq elle est où la pochette cinq elle est là

Mais on ne peut pas savoir

quelque part je ne peux pas savoir non plus je ne peux pas savoir non plus je suis donc quoi je suis dans l'incertain je suis dans l'incertain et dans cet incertain j'ai une décision à prendre alors pour prendre cette décision dans l'incertain qu'est-ce que je suis obligé de faire qu'est-ce que je suis obligé de faire dans mon dans ma décision dans l'incertain

On a la probabilité que la pochette

On va essayer pour le moment on va essayer

86 *On essaie de tout dire sans prononcer le mot probabilité d'accord parce que l'idée qu'on a c'est d'essayer de se dire qu'est-ce que c'est cette probabilité donc on va essayer de le dire sans prononcer le mot probabilité*

On va prendre la boite pour la quelle on a plus de chances de cas

Alors pas de probabilité pas de chances alors vas-y

Le maximum de cas favorables sur le nombre de cas total

D'accord et cet indicateur cet indicateur c'est un indicateur qui va nous permettre de prendre une décision c'est un indicateur qui va nous permettre de prendre une décision maintenant si je prends ce fusible cinq qui nous dit qu'il marche qui nous dit qu'il ne marche pas je ne sais rien du tout je n'en sais rien d'accord bon alors et bien on a cette idée qu'on va bien se mettre

87 *On va bien formaliser au moins oralement alors on va dire que déjà dans le problème que nous intéresse dans le problème que nous intéresse et bien on situe on a à prendre une décision sur un objet quel objet choisir j'ai une décision à prendre sur on objet il est là il est devant moi il faut que je choisisse parmi ces objets il faut que je choisisse parme ces trois montages lequel prendre bon qu'est-ce que je sais de cet objet je ne sais rien de cet objet et pour le prendre je voudrais une décision parce que cet objet aie une certaine caractéristique vous écoutez là donc cet objet il a une certaine caractéristique*

88 *Plutôt cette caractéristique elle est inconnue mais j'ai une décision à prendre par rapport à cette caractéristique donc je ne sais rien individuellement sur cet objet par contre quand je prends un ensemble plus vaste un ensemble auquel appartient cet objet c'est-à-*

dire l'enveloppe et bien là dans cet ensemble et bien j'ai une indication je connais la proportion d'objets qu'ont la caractéristique que je voudrais que mon objet ait donc pour prendre cette décision je ne sais rien directement sur cet objet je ne sais pas s'il y a la caractéristique et ce que je sais qu'il appartient à un ensemble plus vaste et que dans cet ensemble plus vaste je connais la proportion d'objets qu'on la caractéristique que je voudrais à savoir la proportion combien marchent par rapport au total

89 *Bon et bien qu'est-ce que c'est la probabilité c'est simplement par rapport à cette décision de prendre ou pas prendre l'objet il y aura ce que j'appellerai ce qu'on appellera un degré de certitude alors je prends deux cas extrêmes qu'est-ce que çà signifié que la probabilité est égal à zéro qu'est-ce que cela signifié que la probabilité est égal à zéro*

Je n'ai aucune chance

Je n'ai aucune chance çà signifié que je suis certain que cet objet que j'ai devant moi je suis certain que cet objet n'a pas la caractéristique que je voudrais donc dans quel cas j'aurais que la probabilité est égal à zéro bah si dans mon ensemble dans mon ensemble plus vaste je n'ai aucun composant qui ait la caractéristique que je veux çà veut dire que toutes les pièces tous les composants aucun marche

90 *Alors qu'est-ce que çà signifié que la probabilité soir égal à un qu'est-ce que çà signifié que la probabilité pour un objet soit égal à un*

Que j'ai toutes les chances

Çà signifié tout à fait qu'on est certain que l'objet a la caractéristique cherchée et puis quand on fait un calcul de probabilité quand on essaie de calculer une probabilité on essaie de se situer par rapport à ces deux bornes entre zéro et un et bien entre ces deux limites j'essaie de me positionner j'essaie de positionner mon degré de certitude voilà qu'est-ce que c'est quelque part la probabilité alors il y a aussi une chose qu'il faut bien avoir en tête c'est que la probabilité

La probabilité le nombre qui va me permettre de décider quel objet prendre ce nombre ce degré de certitude ce nombre ce degré de certitude il va se faire en proportion à l'ensemble plus grand si je connais la proportion qu'ont la caractéristique c'est en raisonnant sur cet ensemble plus grand que je vais que je vais calculer que je vais déterminer ce degré de certitude alors parmi les parmi les différents groupes parmi les différents groupes après que chaque groupe a exposé son point de vue maintenant est-ce qu'il y a des quels sont les groupes que vous avez l'impression

Qu'ils ton fait pour déterminer quel montage prendre dans quel groupe le raisonnement il est le plus proche de cette idée ou en fait et bien on veut raisonner par exemple le montage a et bien pour savoir s'il est pour savoir si je le prends ou pas par rapport au b par rapport au c et bien je n'ai aucune information sur ce montage en particulier mais je raisonne par rapport à l'ensemble de montages que je peut obtenir est-ce qu'il y a des groupes qu'on raisonné comme çà est-ce qu'il y a des groupes qu'on raisonné comme on vient de faire ici est-ce comme on s'intéresse simplement à un seul élément est-ce que vous avez l'impression d'avoir raisonné comme çà

92 **Un peu quand même**

Un peu quand même mais c'est pas ce que vous nous avez présenté par contre qu'est-ce que toi tu as essayé de faire toi

93 **Non j'ai fait un tableau mais**

Qu'est-ce que tu as essayé de faire quand tu as fait le tableau qu'est-ce que tu as essayé de faire tu as essayé d'obtenir quoi

Bah de trouver la probabilité

Non non on ne parle pas de probabilité

Bah le nombre de chances d'avoir
C'était pas le nombre de chances c'était quand tu fais un tableau c'est pas le nombre de chances que tu obtiennes
Tous les cas
Oui tous les cas
Tu as essayé de chercher le nombre de cas possibles qui correspondaient à tous les cas de ton tableau et puis dans ce tableau trouver quels étaient les cas favorables
Ouais
D'accord bon qui a l'impression d'avoir raisonné de cette manière
Vous avez fait comme çà mais c'est pas comme çà que vous me l'avez présenté et comment vous avez fait pour trouver ces valeurs
94 **Théoriquement le montage B a plus de chances de fonctionner si on avait pas fait tous les tous les cas pour chaque élément et par exemple le montage qu'on trouve qu'a moins de chance c'est le C pour le montage B si on avait où les boîtes du C du montage C**
Vous auriez décidé directement vous auriez dit quoi
95 **Que si on avait les mêmes boites pour les deux**
…c'est physique la c'est le principe électrique des montages en parallèle des montages en série tout à fait je suis d'accord je sais pas si
Dans ce que tu me donnes je ne sais pas si c'est bien compris ce que je disais je vous ai posé une question c'est est-ce que dans votre groupe à un moment vous avez essayé finalement de trouver de raisonner sur tous les montages possibles c'est çà l'idée je ne sais rien sur mon montage en particulier je vais raisonner sur tous les montages possibles je suis ici là le montage a tous les montages possibles et dans tous ces montages possibles quels sont ceux qui marchent j'en ai combien qui marchent je n'ai combien qui ne marchent pas c'est ce que vous aviez fait
Moi j'ai fait comme çà
Nous on avait commencé tu vois avec les oui oui non oui
Alors ce que je vous propose c'est déjà il y en a qui ont fait des calculs pour trouver quelque chose qui permet
96 *Je veux dire sans savoir alors en se referant à des formules anciennes dans savoir de la validité de ces formules s'il faut multiplier s'il faut additionner je vous demande maintenant à tous de revenir vérifier au moins si vos formules sont justes et bien d'essayer pour chaque chacun de ces montages de m'énumérer de m'énumérer tous les cas tous les cas possibles et puis par rapport à tous ces cas quels sont les cas possibles quels sont les cas possibles que nous permettaient de voir ceux qui fonctionnent ici il y en a qui ont réfléchi au problème il y en a qui ont fait une partie mais vous ne l'avez pas fait pour tous les composants*
Oui oui
Précisément j'ai fait tout
97 *Est-ce que votre groupe a dit est-ce que votre groupe a dit vous avez fait quoi vous m'avez présenté un tableau et puis pour trois par contre pour les trois composants vous étiez coincez*
Non non si j'ai réussi
Un tableau avec

trois entrées
Mais non tu ne peux pas

Ah oui ah bon alors vous faites qu'est-ce qu'on a comme technique on a comme technique faire un tableau on a comme technique faire un arbre et puis bah on prend le temps de le faire donc je vous demande maintenant de la faire pour chacun des composants

Pour l'a j'avais trouvé j'ai perdu ma feuille
Fais celui-là d'abord

Commencez par les deux premiers composants montages a et b vous faites bien vos tableaux et regardez si ce que vous avez trouvé çà correspond aux calculs

Vous l'avez fait chacun
Oui
Maintenant vous le faites comment il faut parce qu'il faut laisser les traces comment vous l'avez fait le tableau
C'est-à-dire
Bah c'est-à-dire chercher tous les cas pour trouver le nombre de cas possibles
C'est ce qu'on a fait on a fait les oui et les non
Oui mais pas comme un tableau
Oui mais à peu près la même chose
Pour le montage B alors vas-y explique moi ce que tu as fait
On a fait par rapport à

Là j'ai fait le binaire
Ah oui oui oui non c'est toute la liste
J'ai fait un truc très compliqué moi
Oui
Pour tous les cas six cas oui et trois qui ne marchent pas c'est çà
Mais il y a que quatre cas c'est çà
J'ai pris les
Pour le B eh

	autres tu vois c'est six cas sur neuf six cas sur neuf que çà marche			
	C'est un tableau si tu veux mais plus rapide qu'un tableau		**Bah pour le B il n'y a que quatre cas**	
	Si tu veux		*Ah d'accord la tu raisonnes tu raisonnes çà marche çà marche pas*	
	On va le faire bien là		**Oui**	
			Et moi je veux te dire c'est avec chacun des composants qu'il faut que tu raisonnes tu vas raisonnes là j'ai pour le montage B là dans la deuxième enveloppe j'en ai sept et je peux avoir je peux en avoir deux qui tu travailles sur ces sept composants	
			J'avais bien démarré quand même	*D'accord tu as fais des calculs maintenant on va essayer de comprendre ces calculs*
98		**Dix huit plus neuf çà fait vingt sept**	**Mais c'est quoi N et O ah d'accord c'est non et oui et pour N on a c'est pour le A ou pour le B**	**Mais j'ai fait le tableau déjà c'est le tableau de carnot**
				D'accord alors est-ce que là tu as un tableau que te donne tous les cas possibles par exemple pour moi qu'est-ce que c'est qu'un montage possible dans le montage A un montage possible pour l'A en tout j'en

ai combien de connecteurs j'en ai sept connecteurs je prends un de ces sept connecteurs et puis je prends un fusible et là tu as un montage possible et tu as combien de montages possibles vous pouvez me faire la liste de tous les montages possibles qu'on peut obtenir sachant que tu as sept connecteurs et six qui marchent et un qui ne marche pas donc fais ton tableau

De tous les cas possibles

Tiens regarde en deux colonnes j'ai trouvé la réponse à la base c'est çà et après tu faire la formule et au lieu de mettre à chaque fois oui oui oui faux faux faux comme çà tu fait la formule regarde j'ai un tableau comme çà comme çà

On a six fils qui sont bons six qui sont bons et là trois qui sont pas bons et là tu as donc six fois six trente six six fois trois dix huit après six fois deux douze trois fois deux six et là tu fais çà plus çà plus çà çà c'est le pas bons et çà c'est me bons le deux comme c'est en parallèle il faut faire çà plus çà plus çà et çà c'est les

Il y a combien de montages possibles combien faites le chacun vous raisonnes là-dessus faites votre tableau et puis dans le tableau vous comptez

		bons tiens je veux le mettre à propre attends il faut que je change çà		
		Cinquante sept		**Donc six oui çà c'est non çà c'est non çà c'est non çà c'est oui**
	Toi tu raisonnes çà marche çà marche pas mais moi je te dis pour chacun des composants tu vas raisonner avec chacun des composants tu vas voilà pour le montage b dans la deuxième enveloppe j'en ai sept je peux en avoir je peux en avoir deux non et en fonction de ces composants tu décides			
	Faites le bien pour chacun des circuits			
	Fais le un je fais le deux			*Ok et comment tu vas faire est-ce que vous avez une idée de comment à partir de ce tableau*
				On va faire trois tableaux comme celui-là
	…toi tu as combien de montages possibles tu fais toute la liste de montages possibles sachant que pour chacun de nos circuits	**Cinquante soixante trois ouais soixante trois**		**Non non on fait deux**
100	*Donc tu fais ton tableau fais ton tableau …il y a combien d'ampoules …vous raisonnez là-*	**Alors trente six divisé par deux trois cinquante sept**		**Tu fais AB BC tu fais A plus B je ne sais pas**

	dessus faites chacun à sa manière faites votre tableau et ce tableau vous permet bien voilà exact çà marche c'était çà ton idée en fait dès le début			
	Oui mais en fait	*Alors vous l'avez trouvé pour le trois là*		
	Tu arrives pas à le empatter	**Oui on a trouvé j'ai fait avec deux composants d'abord et après le résultat je l'ai remis avec le troisième composant**		**Je ne sais pas comment faire un tableau pour me troisième comment tu fais pour faire un tableau pour le troisième**
	…	*Et vous ne voyez pas autre moyen de faire*		
101	**J'ai fait avec deux composants d'abord et après le résultat**		**Non**	
		Ils sont en série les composants on fait comme çà parce que s'ils sont en parallèle on ne peut pas comme çà	*Allez je veux les traces pour chacun les traces du tableau maintenant même les brouillons on va les ramasser pour savoir comment vous avez cherché*	
		Mais j'ai fais avec deux composants avec deux composants çà m'a donné trente bons et trente et un pas bons		**Un truc comme çà je crois**
	Tiens tu vois ils ont trente six sur soixante trois	*D'accord tu as cinquante six et là tu en a combien là*		
102	**Ouais çà fait quinze ans que je l'avais dis trente six sur soixante trois**	**Voilà j'ai fait au début huit sept et un pas bons et huit**		
	Donc j'avais bon	*Et avec les deux*		

		premiers composants	
	Cinquante sept et quarante deux ouais c'est çà quarante trois	**On a cinquante six bons et sept pas bons parce qu'il y a avait cinquante six fils bons et sept pas bons quoi et comme çà un tableau**	
103	**C'est pareil pour l'autre regarde**	*Et avec comme çà*	
	Et pour le troisième truc il faut faire quoi trois tableaux	**Et c'est pareil avec l'autre**	
	Ouais	*D'accord*	
	Combien des colonnes		
104	*Mais comment tu as trouvé tout çà*	**Et çà me donne à la fin trois cents quatre vingt …bons et cent quatre vingt quatre pas bons et çà fait**	
	Moi j'ai multiplié deux fois six deux fois cinq	*D'accord vous notez chacun votre truc j'aimerais le fasse sur sa feuille*	
	Et toi tu as fait quoi trois tableau différents	**Tu as fait le c déjà**	
	Oui mais c'est chien	**Non moi je fais d'abord le cas b**	
	Ils l'ont trouvé directement	**Est-ce que tu as fait le cas a est-ce que tu as fait le cas a**	
	Si s'on multiplie	**Oui le cas a je l'ai fait comme**	
	Et toi tu as mis quoi là	**Bah le cas a il l'a fait avec le tableau non**	
108	**Sept un sept un**		
	Sept un sept un		
	Tu as trouvé le pourcentage déjà	**Oui**	**Je n'ai pas compté comme çà**
	Moi j'ai soixante seize pour cent	**Oui mais çà marche pas**	**On deux cinq sept**
		Mais si tu as compris ce qu'on t'as dit	
		Oui quarante deux	*Il faut trouver les*

		et vingt et un	*même résultats dans les calculs tu vois çà çà va nous expliquer* **çà** *va mieux nous expliquer les calculs que vous avez faites*
	Le même truc qu'avant	**Pourquoi tu fait quarante deux regarde là regarde il faut mettre un des deux qui est pas bon et là c'est pareil tu additionnes çà çà et çà et çà donne ceux qui ne sont pas bons**	
	Ouais	**Et pour le c il faut une tableau un gros quoi**	
	Mais deux tableaux	**Là c'est deux cinquièmes et là cinq çà fait quatre non six donc soixante dix donc quatre vingt douze**	**Déjà pour le dernier on peut prendre ce tableau là**
	Trois tableaux	*Et comment tu as fait pour trouver ton trente dix et quatre là*	**Oui**
109	**On a fait deux**	**Bah j'ai multiplié deux fois deux quatre deux fois six douze deux fois cinq dix et cinq fois six trente çà c'est les bons si vous voulez et les pas bons là pareil et là pareil**	**Tu as pas les même valeurs**
	C'est mal arrondi	*Donc là il y a douze montages qui sont pas bons*	**J'ai pas les mêmes oui**
		Ouais	**Ou alors on mets une autre entrée tu fais tac tac tac**
		Et après là dedans tu dis que	
		Ici les bons et là les pas bons il y a	

soixante huit pour cent de bons et trente deux pour cent de pas bons
Mais çà fait combien en total
Tris cent quatre vingt douze bons et cent quatre vingt quatre pas bons
Tu as vu là ils ont fait un tableau et à la fin tu as tout
Je ne sais pas
Donc pour le b tu trouves
Dans le b je trouve dans le b je trouve vingt six bons et trente pas bons
Attends çà et çà c'est les bons et çà
Non çà c'est la dérivation juste les données
Un des deux qui marche là çà marche là çà marche et là çà marche pas
D'accord et le dernier huit et huit
Deux trois quatre cinq six là c'est trente deux çà çà fait trente deux
Regarde lui il marche tout le temps regarde quand on a non et non que çà marche pas tu vois dans le parallèle et c'est deux qui marchent et quatre qui ne marchent pas non non deux oui et cinq non
Non regarde
Ah çà c'était non çà

111 *S'il vous plaît vous m'écoutez maintenant bon pour les cas a et b vous avez trouvé quoi comme proportion combien de bons vous avez trouvé comme proportion combien il y en a de montages qui marchent par rapport au total de montages vous avez trouvé combien j'ai posé une question là*

Moi j'ai fait le tableau j'ai fait juste çà moi

Mais une fois que vous avez faites votre tableau le tableau exploitez-le là Florian tu as fait ton tableau tu as trouvé combien

Non

Bon fais le j'imagine que vous avez de croix ou des bon et mauvais des oui et des non

112 *Après en analysant ce tableau on verra quel est mon degré de certitude on va prendre une décision entre le montage a et le montage b et après on regardera le c*

Çà correspond çà correspond aux calculs que vous aviez fait

Cinquante six pour l'A

Aha et il y en a combien qui marchent fais attention que tu as un montage en parallèle en fait tu as fait comme l'A là

Et comment on fait pour le B

Tu as trouvé combien là

Pour l'A

Pour le B

J'ai trouvé vingt six possibilités sur

114 *Tiens dans le cinq minutes vous avez bien fait pour le cas a et pour le cas b vous avez bien fait y compris pour ceux qu'ils ont fait des calculs autrement Jeremy j'aimerais bien que tu viennes au tableau et que tu nous expliques comment tu as fait pour la cas c vous écoutez tous donc le problème du cas c c'est de trouver tous les cas possibles et parmi tous les cas possibles on a alors vas-y expliques nous écoutez bien tous*

Pour le cas c j'ai pris deux éléments j'ai commencé avec deux éléments huit bon et un cas pas bon et sept bon et un pas bon et çà va donner en fait que cinquante six chances que çà va marcher et seize que çà marchait pas

115 *Tu as cinquante six bons et seize pas bons*

Oui donc pour les fils j'ai trouvé cinquante six bons et cinquante six pas bons et je revient au deuxième au troisième élément plutôt et on fait pareil on trouve les chances çà donne deux cent quatre vingt douze bons et cent quatre vingt quatre pas bons

J'ai trouvé çà intéressant il y avait des points de vue qui s'affrontaient je que je dirais c'est que entre entre qu'ils sont pas tous blanc noir quoi ils ne sont pas tous neuf ils sont entre les *deux quoi donc il y a du savoir ancien il y a un savoir ancien il y a des choses finalement qu'est pas mal quoi est-ce qu'il y a du temps de est-ce que c'était pas si net que çà çà je trouve çà intéressant si tu veux j'ai trouvé intéressant on retrouve ce que je fais souvent quand je fais travail de groupe tu as des points de vue différents qui se confrontent et puis que je pas trop je ne sais pas trop comment çà se confrontent il y a eu un moment quand j'ai demandé bah c'était bête moi je savais pas trop quoi faire j'avais du mal je savais pas trop quoi faire je pensais à la reflexion de ma femme tout à l'heure sur la multiplication on passe un temps pour faire pour tout comprendre et peut être çà va leur fixer des choses ou ce recours scolaire quoi à la formule qu'on a appris quoi un algorithme on connaît les proportions on les multiplies on les additionne et on n'a rien compris est-ce que tu vois donc donc je me dis çà vaut peut être le coût j'étais plutôt satisfait parce que j'ai vu il n'y a eu aucun qui n'est pas rentre dedans quoi moi j'avais peur qu'ils me disent attendez on a fait des calculs quoi c'est bons les calculs on les a fait on les connaît les calculs mais non le fait de le faire pour trouver quelque chose çà leurs à conforté quoi donc je trouve que c'est bon je ne sais pas pour toi dans ta problématique bayesienne fréquentiste est-ce que est-ce que c'est la problématique je ne sais rien du tout çà mais pour moi en termes de pédagogie ancien nouveau (quand tu dis ancien nouveau u veux dire) parce qu'ils ont vu des choses de probabilité et là ils ont des décisions à prendre alors que jamais on leur demande de prendre une décision et on leur dit pas comment faire ou on leur laisse libre moi je crois qu'il y a des choses intéressantes qui se font que je pense que ceux qu'ont pris des formules et bah sachant pas trop laquelle choisir et bah ce recours à bah le fait que tout le groupe ce que j'ai trouvé quand même intéressant ce que m'a un peu fasciné dans les débats c'est que il y a des groupes il y avait des groupes qu'avaient dénombré quoi et çà pouvait servir pour d'autres et dire de le faire et dire tiens est-ce que finalement parmi ces calculs qui je faisais qui apparaissent qui va me permettre de lui donner un peu de sens quoi de lui donner du sens Jeremy je trouve çà excellente ce qu'il a fait c'est vachement bien çà (à moi ce que m'a attiré un peu l'attention j'ai eu l'occasion de parcourir un peu la production de chaque équipe mais d'une part disons que la probabilité est sortie naturellement pour prendre la décision je m'attendais à un petit peu de conflit et non ils ont pris ils l'ont utilisé) ouais et c'est celui qui a produit les choses les plus intéressants bah que je trouve les plus intéressantes et qui disait va savoir (et je me demande si çà été difficile cette manipulation matérielle pour toi de le faire sur place çà a été un peu) tout dans coup je savais pas comment faire la connexion en parallèle et puis je voulais le tester pour voir si tous étaient bien ceux qui marchaient et ceux qui ne marchaient pas je pense qu'à leur niveau ils ont tous bien rentré dans l'activité là je* *les ai eu plus proches ils ont bien rentré que la dernière fois la dernière fois je ne sais pas qu'est-ce qu'ils ont sortie de çà*

Transcription séance La bouteille

Date : 5 mars 2007
N° élèves : 17

Observations :

- Premier colonne : temps [min]
- Enseignant : texte en *italique*
- Elèves : texte en **gras**
- Message adressée à toute la classe : colonne unique
- Message adressée à un ou plusieurs membres de la classe : plusieurs colonnes

Pour l'instant vous n'aurez pas besoin de l'ordinateur on vous dira tout ce qui est matériel on vous donnera des feuilles je vous demande de bien avoir de quoi écrire je vous demande aussi de travailler par binôme travaillez pas à quatre même si vous êtes sur la même table vous éteignez les écrans pas les ordinateurs on va commencer le problème qu'on va vous poser et bien il va nécessiter un petit peu de matériel comme vous voyez alors le matériel que va nous intéresser c'est une bouteille pour le moment il n'y a rien dedans donc c'est une bouteille opaque qui dit opaque dit je ne sais pas je ne sais pas ce qu'elle contienne par contre elle dispose d'un petit bouchon

5 *Ici je pourrais en la retournant connaître le contenu que je vais mettre dedans alors j'ai aussi ici une urne dans laquelle j'ai un certain nombre de boules alors la seule les seules compositions de boules que j'ai des boules qui sont oranges et puis j'ai aussi des boules des boules qui sont noires alors la première chose que je vais faire je vais mélanger on connaît pas je précise je précise on connaît pas la proportion qu'il y a de boules noires et de boules oranges et bien ce que je vais faire c'est que je prend quatre boules au hasard*

6 *Je vais les mettre dedans quatre boules et le but du problème çà va être de voir une méthode qui nous permettrait d'estimer de voir comment on pourrait estimer la composition de cette bouteille opaque alors j'en prends quatre bon alors çà c'est la première chose j'ai cette j'ai donc cette bouteille opaque elle contient elle contient quatre quatre billes quatre boules et il s'agit de trouver une méthode qui va nous permettre*

7 *D'estimer la composition la composition de cette bouteille alors pour ce pour ce faire la première question qu'on pourrait quand même se poser c'est cette bouteille qui contienne quatre boules quelles seraient quelles sont les différentes compositions possibles alors je peux avoir je peux avoir quatre noires je peux avoir quatre oranges je peux avoir*

Quatre noires quatre oranges
Deux noires deux oranges
Une noire trois oranges
Trois noires une orange

Je peux avoir une noire trois oranges je peux avoir deux oranges deux noires je peux avoir trois noires une orange

Mais on ne peut pas savoir

Mais je te dis on cherche les compositions possibles d'accord

Ok les compositions possibles

Donc comme compositions possibles on a combien en tout

8 **Il y en a cinq**
Il y en a cinq d'accord toutes les boules noires une orange quatre noires deux oranges et deux noires etc. etc. donc çà c'est la première chose donc voilà je le note de cette manière pour vous aider je vais dire que j'ai les cinq compositions possibles et puis je vais leur donner un nom toutes les quatre noires sera A etc. etc. on les nombrera dans la suite A B C D et E autre chose qu'on va faire pour essayer de d'estimer pour trouver une méthode qui va nous permettre d'estimer la composition de cette bouteille

9 *c'est de travailler avec des pièces de monnaie donc je vous donne à chaque groupe je vous donne une enveloppe qui contient vingt pièces de monnaie*
Pour quoi les monnaies
Contrôlez que vous en avez bien vingt la deuxième chose que je vais vous donner pour travailler

10 *Commencez pas à jouer avec s'il vous plait donc la deuxième chose que je vous donne c'est une feuille que va nous servir de base pour travailler donc je vous donne ces feuilles alors je commente le tableau sur lequel vous allez travailler donc on a une situation au départ notre situation au départ c'est exactement celle-là*

11 *vous avez sur le vidéo projecteur la bouteille on ne sait rien de sa composition on ne sait rien de sa composition mais on sait qu'il peut y avoir cinq compositions possibles la règle du jeu c'est qu'on va essayer au fur et à mesure de compléter ce tableau vous allez compléter ce tableau au fur et à mesure alors comment est-ce que vous allez faire vous avez là les consignez je ne les donnerai qu'oralement le premier principe c'est déjà pour bien comprendre le tableau donc la première colonne du tableau on indiquera O ou N orange ou noire la couleur de la bille que l'on observera à travers le bouchon lorsque moi je retournerai la bouteille quand je retournerai la bouteille vous noterez dans la colonne de gauche la première colonne*

12 *Vous noterez ici dans la colonne de gauche vous noterez et bien N ou O donc çà c'est la première chose deuxième chose et bien et bien dans le zone grise du tableau vous laisserez les traces du nombre de pièces que vous attribuez à chaque composition alors je vais m'expliquer un petit peu plus en détail un petit peu plus en détail voilà la règle la règle qu'on va adopter c'est que la répartition que vous allez à repartir vos pièces de monnaie affecter des pièces de monnaie à chacune de compositions alors la règle c'est la suivante*

13 *Bah les règles c'est que plus vous croyez en une composition et plus vous allez lui attribuer de pièces bien sûr toutefois si vous croyez plus en une composition qu'en une autre vous essayerez aussi de justifier c'est-à-dire que le travail est à la fois d'estimer suivant votre degré de croyance de telle ou telle composition de l'estimer avec ces pièces et puis au même temps et bien on vous demande aussi je vous demande d'essayer de justifier vos choix pourquoi pourquoi est-ce que vous attribuez tel ou tel nombre de pièces alors comme c'est un travail à deux il faut aussi que vous vous mettez d'accord et quand vous êtes d'accord sur la distribution de pièces et bien vous les écrivez*

14 *Vous justifiez dans les casses respectives le nombre de pièces que vous attribuez à chaque composition je vais continuer encore pour vous pour compléter ces règles un exemple un exemple si vous lui attribuez aucune pièce monnaie à une certaine composition par exemple je prends la composition C c'est-à-dire NNOO si vous lui attribuez aucune pièce de monnaie çà signifié que vous êtes sûrs que la bouteille que la composition de la bouteille n'est pas du type C c'est-à-dire NNOO et puis de la même manière si vous attribuez toutes les pièces à la composition E c'est parce que vous êtes sûrs que la bouteille est du type OOOO voilà les règles alors et bien*

15 *Voilà les règles dans la première colonne ma situation de départ c'est celle-là je vous demande alors j'ai pas encore touché la bouteille je ne sais pas ce qu'il y a dedans vous*

non plus je vous demande d'attribuer dans la première ligne donc dans la première ligne vous attribuez le nombre de pièces que vous pensez pour chacune de compositions possibles ok allez vous faites çà donc je vous le redis faites le à deux faites le pas à quatre çà perd du sens si vous le faites à quatre

16 *Avant d'écrire il faut vous mettre d'accord*

17

Pareil vous vous travaillez ensemble là première ligne la première ligne la deuxième ligne c'est pour quand on prend une bille		**C'est quoi les règles j'ai pas écouté les règles c'est quoi là**	Ne sachant pas combien il y a de boules noires ni de boules oranges le truc est équiprobable
On mets dix au milieu et cinq aux autres		**Je ne sais pas**	
Alors vous essayez de justifier pourquoi vous mettez ces pièces là		**C'est marqué là**	*Je suis d'accord vous écrivez votre truc*
Et pourquoi on ne met une là au cas où		**Çà mieux là**	
Il y a plus de chances d'avoir NNOO la C quoi			**Quoi**
Mais çà peut arriver quand même les autres			**Toutes les compositions sont équiprobables**
Wé wé			**On voit pas la différence entre entre toutes les possibilités**
Moi je mettrais dix à la C et pour les autres le reste cinq et cinq		**Moi je mettrais toutes au milieu probablement**	**Wé je suis d'accord**
Oui tu dis au milieu		**Probablement**	
Oui il y a autant de noires que d'oranges		**Je les mettrais toutes au milieu je ne sais pas**	
Mais je mettrais un quand même çà peut sortir quand même		**Moi non plus**	
Mais regarde celle qu'il y a caremment plus de chances d'avoir celle-ci		*Il faut que tu repartisses tes pièces de monnaie*	

	Oui mais là tu l'écartes quoi	**Là c'est complètement aléatoire**	
	Oui mais çà tu imagines	*Tu mets de pièces que correspondent à ton degré de croyance*	
	Çà c'est comme le lotto tu prends de risques ok je vais concéder	**Mais là c'est complètement aléatoire**	
	Alors est-ce que vous pouvez me donner vos explications	**Si tu penses que celle-là est pas du tout possible et bien tu mais aucune pièce et si tu penses qu'elle est possible tu mets de pièces**	
	Bah c'est plus dur d'avoir quatre boules de la même couleur	**Mais monsieur c'est complètement aléatoire**	
	En fait monsieur il faut faire trois fois comme çà	**Moi je mettrais toutes au milieu là**	
	Non non pour le moment c'est la première ligne juste la première ligne la deuxième ligne c'est quand tu prends une boule d'accord essayer de l'expliquer maintenant	**C'est la probabilité il y a vingt pièces il faut voir par rapport à la statistique par rapport à la statistique**	
	Si il y a autant d'oranges et de noires	**Attends on fait vingt divisé par cinq**	
	Il est rare d'avoir toutes les boules identiques	**Là huit pièces un deux trois quatre cinq six sept**	
	Oui il y a carement plus de chances d'avoir la C que l'A	**Tu mets huit au milieu**	**Si on fait un truc d'égalité c'est simple quoi**
18	*Le suspense monte*		
	Alors est-ce qu'il y a forcement autant de boules noires que d'oranges	**Après on va mettre quatre là quatre là et après deux et deux**	
19	*Bon c'est bon tout le monde vous avez terminé*		

Le suspense devient insoutenable

20 *Bon alors je vais retourner la bouteille donc voilà je retourne la bouteille et j'obtiens j'obtiens orange je vais la laisser comme çà donc çà veut dire çà veut donc dire que maintenant vous avez une bille ici elle est orange*

21 *donc vous savez que cette bille est orange vous allez remplir le tableau donc vous vous mettez un O là vous mettez O ici donc vous avez déjà une bille et maintenant maintenant que vous savez que cette bille est orange et bien avec cette information nouvelle que vous arrive voyez si la croyance que vous aviez est-ce qu'elle évolue est-ce qu'elle va changer est-ce que vous allez repartir autrement vos pièces d'accord*

22	**Non pareil bah non attends çà çà veut dire que là maintenant on mets zéro**	**On barre A**	Je vous propose de réfléchir déjà comment on ferait si après il tire une orange ou une noire sachant que il va nous demander s'il a tiré une ou l'autre une orange ou une noire
	On devrait pas avoir mise comme çà tu vois	On mets un quatre dix neuf	S'il tire une autre orange on sera plus équiprobable on sera plus proche d'oranges s'il tira une autre orange on n'aura ni la A ni la B comme possibilité
	Là il y a plus de probabilité là je mettrais une pièce quoi	Je ne change comme çà	Pas forcement parce que çà peut être la même boule orange là çà serait une chance sur quatre que ce soit la même
	Là disons il y a autant de probabilité	**Quatre quatre et là on mets plus**	
	Vas-y comment	**Je dis que B que là on mets plus**	
	Çà serait zéro quatre sept cinq quatre	**Mais tu fais comment ton quatre quatre là**	
	Wé çà fait combien vingt	**Tu as mis combien là**	
	Combien sur l'A zéro et pour les autres de toutes façon c'est pareil on sait pas		

23 *Donc ce que m'importe aussi ce que m'importe ce que m'importe c'est que vous arrivez à justifier pourquoi vous les faites évoluer pourquoi au départ vous avez mis tel ou tel*

	nombre de pièces et pourquoi est-ce que çà évolue pourquoi est-ce que vous le faites évoluer la distribution de pièces essayez d'expliquer pourquoi vous avez fait çà		
	Non non pas nous on a divisé la thune	**Et pourquoi**	
	Il faut voir après quatre boules	**Bah parce que après il peut soirtir une noire**	
		Il faut justifier là	
		On met combien de pièces là	
		Un deux trois quatre	
		Et six çà fait quatorze	
24	*Et je vous demande aussi une chose tout ce que vous écrivez surtout ne le gommez pas si çà évolue si vos idées évoluent au fur et à mesure vous le noterez mais ne gommez surtout pas vos premières idées bon expliques nous Florian mettez vous d'accord*		
	Il y a plus de chances de tirer celle-ci que celle-ci	**Là il y a plus de chances d'avoir il y a une chance sur deux**	**Ici là il faut barrer çà**
	Mais c'est quoi le but de tout çà	**Le nombre est en fonction du nombre de pièces oranges**	**oui**
		En fait il faut appliquer si on a autant de chances de noires que d'oranges	**La question est comment décaler tout çà par là**
		Une chance sur deux	**Parce que même si c'est possible que se soit la même boule orange et que aucune des trois boules noires n'est sorti il y a une possibilité quand même parce qu'il y a une chance sur quatre pour la même bille**
25	*Bon je continu donc j'ai fait le premier retournement çà y est et maintenant je vais effectuer le deuxième retournement je mélange et j'obtiens noire*		
26	**Donc là on a noire**	**Donc là il y a zéro**	**Maintenant**
	Ecoutes on fait pareil on mets dix cinq et cinq	**Oui çà on peut la barrer**	**Noire et orange çà serait plutôt celle-là on a eu une noire et une orange**
	Là il faut faire un celui-là deux celui-	**Parce que on a aucune chance**	Wé

là et un celui-là çà fait huit six six	**d'avoir toutes oranges**	
Çà va pas		**On élimine ces deux cas seulement**
Il faut voir la probabilité		**Wé**
Justement il faut que soit pareil		**Donc il faudrait voir une troisième**
Là regarde tu mets six et six quatorze douze douze et huit çà fait vingt		**Et comme oui oui juste une troisième**
Tu prends deux là et deux là		**Monsieur on est obligé de mettre les vingt pièces**
Tu mets pas un là		*Wé wé wé*
Si si on met un là là j'en mets deux pour être à six et il me reste un pour là non non un un seul		**Moi je ferais comme çà**
Oui mais tu peux mettre dix cinq cinq		**Donc si çà était équiprobable et là il faut choisir comme on a eu une noire et une orange et que là il y a autant d'oranges que de noires on mets dans le C**
Oui c'est pareil		**On met deux**
Je dis que maintenant il y a plus de chances d'une noire		**Donc çà fait six huit six et zéro**
		Wé
		Donc zéro six huit six zéro et çà fait vingt

27 *Bon alors ce que je vous propose maintenant alors que vous avez on va essayer de mettre en commun tout çà je vous demande de bien écouter ce que chaque groupe dit intervenir si vous êtes d'accord si vous êtes pas d'accord mais ne pas intervenir tous ensemble sinon on ne pourra pas s'écouter*

28 *On est on va prendre étape par étape donc la première étape donc pour le moment on n'a rien tiré qu'est-ce que vous avez répondu comment vous avez reparti vos pièces et pourquoi alors je vais demander à Jenny déjà nous dire ce que tu as réponde*

Pour le premier

Avant de faire le premier disons la situation de départ

On a mis deux quatre huit quatre deux

29 *Bon alors est-ce que vous justifier là pourquoi vous avez mis çà*

On a autant de chances d'avoir une boule orange et une boule noire on peut pas le savoir donc on a une chance sur deux

Vous avez autant de chances d'avoir une boule orange et une boule noire alors pourquoi est-ce que vous faites deux quatre huit quatre deux Jenny tu as une explication

Bah comme on a cinquante pour cent de chances d'avoir des boules oranges et des boules noires donc celui pour lequel on va miser plus c'est deux boules noires et deux boules oranges et après on a moins de chances pour les autres

Oui mais théoriquement on a la même la même chance pour les cinq possibilités au départ

D'accord est-ce qu'il y a des réactions j'entendu des élèves qui voulaient intervenir

Je ne sais pas s'il y a cinquante pour cent de chances de tomber sur une boule noire et d'orange

Wé

Wé

En fait on ne connaît le contenu de l'urne

Çà on ne le connaît pas

30 *Bon alors quelqu'un veut intervenir*

Wé théoriquement on a autant de chances au départ d'avoir les cinq possibilités

On a autant de chances d'avoir les cinq possibilités est-ce que c'est çà que vous avez répondu

Wé

Donc vous avez mis quoi

Quatre quatre quatre quatre quatre

Vous avez mis quatre quatre quatre quatre et quatre qui d'autre a mis ceci et qu'est-ce que vous donneriez comme justification qu'est-ce que vous avez donné comme justification

Bah qu'il y a autant de chances de probabilité de tomber sur ces possibilités

Est-ce que parmi alors il y a deux groupes il y a vous êtes deux groupes d'avoir répondu çà les autres est-ce qu'il y a d'autres qu'on répondu çà et les autres qu'ont répondu autrement qu'est-ce que vous diriez

31 *A ces deux groupes est-ce que çà paraît convainquant ce qu'ils disent est-ce que çà quel est votre argument pourquoi vous avez répondu autre chose çà vous parait peu convainquant tiens toi explique nous*

Nous on a répondu trois six deux six trois

Trois six deux six trois alors est-ce que vous pouvez nous

Bah parce qu'en fait au départ pour les cas de billes …on ne sait pas s'il y a autant d'oranges que de noires et après pour que la situation A et la situation E sont des cas d'avoir pas beaucoup de chances en fait parce que ce sont de billes de la même couleur et après pour la situation C en fait c'est cinquante cinquante

32 *Donc excuse moi je te coupe la parole çà et çà tu m'affectes la même chose ici et ici tu affectes la même chose et puis celle-là est à part alors pour quoi est-ce que vous avez mis d'autres en fait on voit que si on a répondu à cette partie du tableau on a l'autre maintenant pourquoi est-ce que vous avez répondu trois six deux pour les trois premières lignes*

….on sait que…on sait qu'on peut avoir deux chances on ne connaît pas la proportionnalité du début d'oranges et de noires donc est …mais on sait que c'est rare de tomber sur quatre noires ou sur quatre oranges donc pour cela on va

commencer par mettre
Pour cela vous avez mis çà à part
Mais justement s'il y a que de noires tu auras NNNN toujours on ne sait pas
On ne connaît pas trop la proportionnalité et on a dit de partager quoi
Alors tu ne connais pas trop la proportionnalité tu ne connais pas la proportion et tu dis que celle-là elle va avoir moins de chances que celle-là pourquoi
Parce qu'en fait si on avait le même nombre de si on savait par exemple qu'il y a avait deux billes oranges et deux bille noires des le départ on aurait plus de chances d'avoir ce cas-là…mais on ne sait pas et il peut y avoir trois billes oranges et une bille noire
33 *Et pourquoi est-ce que le fait d'avoir trois billes oranges c'est ce que vous me dites pourquoi le fait d'avoir trois billes oranges et une bille noire çà vous semble plus probable que d'avoir deux noires et deux oranges c'est ce que vous mettez je vous parle c'est ce que vous mettez comment est-ce que vous justifiez çà est-ce que vous le justifiez est-ce que çà vous parait cohérent est-ce que çà vous parait cohérent*
Mais on ne peut pas dire qu'il y a cinquante pour cent de noires et cinquante pour cent d'oranges
34 *Alors ceux qui ont donné une autre réponse qu'est-ce que vous leur diriez simplement pour dire que vous êtes d'accord ou que vous êtes pas d'accord*
Là par rapport à comme ce sont reparti les pièces je dirai qu'il y a autant de boules noires que d'orange parce que par rapport aux tirages si sur tout les tirages on avait dix tirages on avait trois tirages A six tirages B et deux tirages C six tirages D et trois tirages E çà peut à peu près dire qu'on a cinquante et cinquante

Il y en a qui n'écoutent pas
35 **En voyant par exemple si on faisait des tirages et on a trois tirages A six tirages B deux C six D et trois E en voyant çà je dirait qu'il y a autant de boules noires que d'oranges à peu près cinquante cinquante à peu près**
Wé et puis
Et puis
Normalement çà veut dire que la moyenne çà devrait se rapprocher plutôt de C que de D ou E si c'est à peu près cinquante cinquante
Si on aurait cinquante cinquante
çà devrait se rapprocher de C
çà devrait se rapprocher de C est-ce que c'est le cas là
Bah non
Non
Maintenant est-ce qu'on sait qu'on a cinquante cinquante
Non il peut y avoir quatre vingt pour cent de boules noires
On sait pas alors avec tout ce qu'on a dit avec tout ce qu'on a dit vous vous avez mis quoi
Trois quatre six quatre trois
36 *Alors est-ce que trois quatre six quatre trois c'est-ce que vous avez mis est-ce que vous le changeriez ou est-ce que vous le laisse*
Non c'est bien on le garde
Vous le gardez est-ce que vous pouvez justifier pourquoi vous le gardez
Parce qu'on ne sait pas la proportion des couleurs il y a donc autant de chances d'avoir de chaque de chaque couleur

Je sais que vous êtes pas tous d'accord pouvez vous donner des arguments donc tu me dis que pour toi de toute façon avoir deux noires et deux oranges c'est ce qui doit apparaître qu'il y a plus de chances d'obtenir çà que de ne pas obtenir çà alors est-ce que vous êtes tous d'accord et sûr tout si vous êtes pas d'accord essayez d'argumenter

37 **C'est pareil on ne connaît pas les couleurs qu'on a dans la bouteille**

C'est vrai ce le même problème on ne sait pas le contenu de la bouteille çà c'est sur mais dans ce qu'il dit il me dit qu'il y a plus de chances d'avoir dans la bouteille noire noire orange orange que d'avoir noire noire noire noire sans avoir aucune information sur ce que j'ai mis dans cette bouteille sans avoir aucune information sur la composition de mon urne de laquelle j'ai pris les quatre boules est-ce que vous êtes d'accord avec çà

Comme on ne sait rien il peut il avoir dix pour cent de noires et quatre vingt dix pour cent d'orange pour moi c'est toutes équiprobables

38 *Donc tu nous dis imaginons que dans l'urne il y a quatre vingt dix pour cent là dedans eh imaginez que quatre vingt dix pour cent d'oranges et dix pour cent de noires beaucoup d'oranges et très peu de noires est-ce que vous auriez donné cette réponse*

Non

Est-ce que cette réponse elle convienne si vous aviez çà

Non

Bon d'accord bon alors le problème c'est qu'on ne sait rien et quand on ne sait rien qu'est-ce qu'on a envie de faire quand on ne sait rien bah c'est ne pas trop se prononcer donc pour ne pas se prononcer au départ comme on ne sait rien qu'est-ce qu'on fait on admet qu'on ne sait rien et si on admet qu'on ne sait rien pourquoi est-ce qu'on privilégieriez comme vous l'avez fait pour beaucoup pourquoi est-ce que vous privilégierez une situation plutôt qu'un autre pourquoi pourquoi est-ce que tu me dit tiens le C je vais l'avoir plus souvent c'est parce que tu dis

39 *Tu le dis parce que tu fais une supposition une supposition de la composition de départ mais si tu ne sait rien de cette composition autrement dit tu peux pas le justifier ton six et je ne sais pas si vous l'avez constaté mais dans ce que j'avais marqué on trouvait exactement le contraire au niveau ordre il y a en qui m'avait mis deux ici six ici trois et six et trois bon alors çà correspond à quoi ce deux réponses çà correspond à quoi ces deux réponses çà correspond à un a priori*

40 *que vous faites par rapport par rapport à l'urne qui m'a servi mais moi mon problème c'est que je ne sais rien sur cette urne donc quelque part je ne pas prendre position je ne peux pas prendre position et donc si je ne peux pas prendre position je ne peux pas privilégier a priori comme je ne sais rien a priori plus une composition qu'une autre je ne sais strictement rien et puis comme je ne sais strictement rien et bien qu'est-ce qu'on a comme réponse*

Quatre quatre quatre quatre et quatre

N'effacez rien de ce que vous avez mis bon on continu maintenant on va partir de quatre quatre quatre quatre alors vous me dites non non ne le changez pas donc vous allez me le rappeler c'était orange la première bon alors avec cette nouvelle donnée alors

On peu barrer NNNN

Après on mets cinq cinq cinq cinq

On peut barrer autrement dit on peut mettre zéro ici alors qu'est-ce que vous avez fait et comment vous le justifierez

Cinq cinq cinq cinq

Cinq par tout

Tu mets cinq par tout

41 *Donc tu mets cinq par tout pourquoi*

Bah pareil on ne connaît la mise du début

Wé il a raison
Parce qu'on ne sait pas on ne sait pas que qu'il y a dedans tout le monde est d'accord tout le monde est d'accord alors je continu la suivante c'était quoi
Noire
C'était noire alors qu'est-ce qu'on va faire
On barre la E
Zéro à la OOOO
Déjà on mets zéro ici
Et après six virgule six à chacune
42 *Là tu essayes de repartir le même nombre de pièces aux autres je vas l'écrire sous la forme de fraction momentanément donc là vous faites vous en avez vingt pièces vous feriez vingt tiers vingt tiers vingt tiers etc. etc. c'et bien çà bon alors réfléchissons un petit peu réfléchissons un petit peu vous avez tiré une fois orange et vous avez tiré une fois noire et vous me dites que ayant tiré une fois orange et une fois noire vous avez la même degré de certitude d'obtenir celle-là que celle-là que celle-là est-ce que vous êtes d'accord avec çà*
En fait comme on a tiré une fois orange et une fois noire on pourrait dire que çà peut être cinquante pour cent qu'il y a des boules oranges et de boules noires dans ce cas on favorise la composition C alors
Toi tu as donc plutôt envie de favoriser la composition C
43 **C'est-à-dire que là on a tiré une orange et une noire mais si on avait tiré une orange et après une orange çà aurait pu être la même bille**
Est-ce que finalement c'est çà …tu aurais fait quoi
J'aurai plus penché pour le B et D
Tu aurais plus pensé pour B et D
B et D
Tu aurais plus mis ici et là et puis moins ici çà alors tu l'aurais repartie tu aurais changé çà tu dis tiens je suis plus convaincu de çà plus moins autant
Autant
Autant et ici tu mettrais plus bon alors je mets plus là je mets sept là … je rajoute la moitié tu veux que je coupe une pièce en deux c'est çà que tu veux faire
Oui
44 *Est-ce que çà est-ce que çà vous paraît convainquant*
Non
Non on ne sait toujours rien
Alors pourquoi Jenny
Est-ce que parce qu'on a tiré une orange et une noire donc çà fait cinquante et cinquante on va favoriser plutôt celle là du milieu
Tu favoriserais plutôt celle du milieu voyez on est pas d'accord on a du mal à se mettre d'accord parce que toi tu me dis je favoriserais plutôt celle là que j'ai rencontré celle-là çà fait zéro mais celle là que je veux augmenter et celle-là je veux diminuer bon on a du mal à se mettre d'accord je vous rassure c'est normal oui tu voulais dire
45 **Je favoriserais aussi la C parce vue que les tirages ont donné autant d'oranges que de noires même si on a fait deux tirages même si comment si la D soit équiprobable et qu'on veut mettre six par tout moi j'ai mis sur la C**
Je dirais que plutôt cinquante pour cent d'orange et cinquante pour cent de noire
Donc tu mettrais quoi là zéro zéro et puis
Six huit six
Six huit six autrement dit là tu dis tu dis je fais évoluer la probabilité bon alors on va essayer de voir si quand même parce que çà vous ne l'avait déjà dit on va imaginer on va imaginer qu'on obtient si on avait obtenu deux boules oranges est-ce que çà aurait changé

votre manière de raisonner et en quoi çà aurait changé votre manière de raisonner

46 Bah on ne peut pas savoir si la deuxième boule tirée orange n'est pas la première

Ah çà oui on ne peut pas savoir on ne peut pas savoir si tu le savais tu aurais pas de problème mais nous on tire on tire deux fois orange alors qu'est-ce que vous auriez envie de faire de rien faire

De favoriser D et E de baisser B et C…

Donc vous auriez envie de plutôt favoriser bah c'est ce que toi tu dis Jenny de favoriser B E

Non de baisser B et d'augmenter C bah C D E quoi

C D E et

Et de baisser un petit peu celle-là

Et de baisser un petit peu celle-là tu peux me dire pourquoi

47 Parce que s'il y a deux boules oranges différentes qu'on a eu B il vaut normalement zéro non mais on n'est pas sûr que ces soient des différentes parce que peut être la même donc B est çà peu arriver quand même on laisse un peu à B

Wé je comprends ce que tu veux dire est-ce qu'il y a d'autres argumentaires à faire Florian tu en penses quoi

La même chose que Jenny la même chose que Jenny

La même chose que Jenny bon alors suivant ce principe tu me dis là si j'obtiens orange toi tu me dis que tu as plus de chance d'avoir çà me fais augmenter mon je reviens à la première étape je viens d'avoir orange ici j'obtiens orange pourquoi on changerait le procédé tu obtiens orange en premier qu'est-ce que tu as envie de dire

48 Bah qu'on favorise l'orange

Qu'on favorise les cas on aurait envie de favoriser les cas où quand ce cas est réalisé on aurait beaucoup de chances d'obtenir orange c'est clair ce que je dis je le répète je le dis on aurait envie de se dire si bien qu'on ne l'a pas eu on ne l'a eu du tout au début d'avoir deux oranges mais si je vois déjà orange apparaître bien sûr que je peux me dire tout de suite j'ai orange qu'apparaît ops je supprime celui-là

Non non

49 *Pardon pardon je supprime A c'est pas possible mais si j'obtiens orange je peux me dire imaginons que soit çà que j'ai dedans imaginons que ce soit réellement OOOO que j'ai dedans eh bien le fait d'avoir orange ici çà a plutôt tendance à me conforter dans l'idée que c'est çà parce que si c'était çà en effet j'aurais j'obtiens orange donc çà j'ai envie de me dire que çà çà pourrait monter je prends ton idée je prends ton idée mais je continu tu continu Jenny*

Bah là c'est pareil çà va augmenter

Celui-là on pourrait dire qu'il augmente en suite

Et C pareil et B un peu moins B devrait diminuer un petit peu

Est-ce que ayant cette ayant cette imaginons que ce soit NNNO trois noires et une orange est-ce que c'est çà que vous attendriez d'avoir

Bah non non non non une noire

Qu'est-ce que vous attendriez d'avoir plutôt en premier

Noire

Une noire

D'avoir une noire donc si vous obtenez orange est-ce que çà vous confirme dans votre idée imaginons que c'est çà

Bah non non non

50 *Non çà fait plutôt tendance à la faire baisser d'accord bon ici*

Il y a assez constant plus ou moins

Tu me dis

Assez constant plus ou moins

Bon pourquoi pas je vais mettre comme çà alors qu'est-ce qu'on pourrait imaginer comme vous avez vingt pièces maintenant vous me dites une idée
Là quatre
Ici je mettrai quatre
Deux ou quatre
Oui deux ou quatre et six six
Deux
Sept et sept si si
Non six huit
Çà fait dix huit
Alors est-ce que le total il fait bien
Oui vingt
Et bien çà marche bon le total il fait bien vingt et est-ce que çà tient en compte les remarques qu'on a fait
wé
çà tient compte les remarques qu'on a fait bon alors on prend et

51 *Et puis on va prendre en compte ce qui c'est passé et voilà maintenant on a une boule noire alors avec cette idée qu'on peut raisonner disons bon voilà j'imagine la composition est-ce que çà c'est par rapport si ma composition était par exemple A est-ce que çà serait possible que j'obtienne noire ici et n'oublions pas que j'ai obtenu une orange çà c'est pas possible bon*
Et E aussi c'est zéro
et puis et là je peux l'enlever c'est physique c'est même pas un problème de probabilité maintenant pour les autres qu'est-ce que vous diriez
D'abord je ne suis pas sûr
Tu mettrais
…
C'est c e que tu avais mis avant mais est-ce que cela tient compte ce qu'on a dit avant

52 *Donc en prenant compte de çà est-ce que çà te permettrait six huit six envie de mettre autre chose bon c'est pas facile bon c'est pas facile de savoir mais est-ce qu'il y a une chose que vous auriez envie de dire il y en a qui devrait augmenter il y en a qui devrait baisser*
Bah déjà C devrait augmenter bien parce qu'on a eu une boule de chaque et B aussi un peu quand même
B D aussi
B B et D on le laisse constant là wé
Oui B était très faible wé
D la laisser constant ou baisser un peu

53 *D'accord on peut hésiter c'est bon déjà on a pas mal avancé on a pas mal avancé dans la perception du problème on a pas mal avancé alors pour essayer d'aller un petit peu de continuer on va garder en mémoire les données et puis vous avez parlé de probabilité c'est-à-dire au moins dans notre manière de raisonner notre manière de dire le choses qu'est-ce qu'il passerait si on parlait de probabilité alors on parle de probabilité je suis dans le premier je suis dans le cas que je ne sais rien qu'est-ce que vous diriez*

54 **Une chance sur cinq**
Vous ne savez rien vous ne savez rien vous avez privilégié je répète l'idée que certains ont dit qu'ont du mal à se convaincre on ne va privilégier une en opposition plutôt qu'une autre je ne sais rien du tout donc on va mettre le compteur à zéro et puis se dire j'ai pas d'information je n'ai pas de raison que je privilégie une donc on met tout à un cinquième bon après qu'est-ce qui se passe alors si je suis là j'ai pas encore retourné ma j'ai pas

encore retourné mon mais a priori je me dis que j'ai autant de chances que là dedans y aie une de ces cinq compositions

55 *Bon alors je ne sais pas ce que je vais obtenir mais nous qu'est-ce qu'on a obtenu on a obtenu*

Orange

On a obtenu orange est-ce que vous pouvez me dire alors déjà cette probabilité comme est-ce que nous on le noterait çà correspond à quoi cette probabilité on aurait écrit P de A c'est un cinquième que P de B c'est un cinquième etc etc donc je résume donc avant le premier retournement ce qu'on a dit on ne sait rien on va pas privilégier une composition plutôt qu'une autre don finalement çà va être un cinquième à chacune des probabilités

56 *Et on a dit que P de A c'est un cinquième et il y a la P de B la P de C il y a P de D la P de E donc alors voilà donc voilà donc ce qu'on a dans notre tableau si je raisonne en probabilité qu'est-ce que je cherche en fait en fin avant le premier retournement avant d'avoir fait le premier retournement voilà simplement ce que je sais mais il y a une question que je pourrais me poser c'est puisque a priori je ne sais que çà est-ce qu'on ne pourrait pas trouver sachant çà partant de ce principe P de A est égal à P de B égal à P de C toutes égal à un cinquième est-ce qu'on pourrait trouver la probabilité d'obtenir O*

57 **Bah non non non**

Cinquante pour cent bah un demi quoi

Alors je vais aller un petit peu plus loin voilà chacune des cinq compositions possibles donc on a chacune des cinq compositions possibles je sais que j'ai une de ces cinq quelle est la probabilité ce que j'aimerais trouver c'est la probabilité alors ne sachant rien la probabilité bah sachant qu'a priori chacune de ces compositions aura autant de chances qu'une autre de sortir la question que je me pose c'est quelle est la probabilité d'obtenir orange bon alors si c'était cette combinaison si c'était cette composition

58 *Çà serait quoi la probabilité d'obtenir orange*

Zéro zéro pour cent

Elle vaudrait zéro celle-là un quart et celle-là

Un demi

Un demi un quart en zéro donc qu'est-ce que c'est cette probabilité que vous me dites égal à un quart qui correspond à B comme est-ce qu'elle se dirait dans notre langage de probabilité un quart çà serait çà serait quoi çà serait la probabilité de quoi çà serait la probabilité d'obtenir orange sachant sachant qu'on a sachant qu'on a B sachant que la composition est B donc quand vous me dites en observant les différentes combinaisons possibles

59 *Que l'on tient chacun de ces nombres avec notre notation qu'on utilise depuis quelques semaines comment çà s'écrirait celle-là çà c'est quoi çà serait la probabilité d'obtenir orange sachant qu'on a la composition A comment tu écrirais çà*

P de

P de

N sachant O

O d'orange sachant

N

Sachant A qu'est-ce que c'est l'A ici c'est la probabilité si on sait qu'on a cette composition-là sachant que la composition de la bouteille est celle-là c'est la probabilité d'obtenir orange

60 *Donc çà c'est le P de O sachant B et çà c'est le P de O sachant C çà c'est le P de O sachant D et çà c'est le P de O sachant E nous dans la question que je vous pose c'est est-*

ce qu'on pourrait trouver la probabilité d'obtenir orange sachant qu'on ne sait pas qu'on ne peut pas privilégier une de ces compositions plus qu'une autre comment est-ce qu'on pourrait faire comment est-ce qu'on pourrait faire pour calculer cette probabilité d'avoir orange alors des idées je m'appui sur des techniques qu'on a mis en place qu'est-ce que j'ai comme compositions possibles A B C D

Et E

61 *Et moi je m'intéresse à quoi et bien je m'intéresse à la probabilité de la boule que je tire quand j'effectue un retournement est-ce qu'elle sera orange ou noire saisissez qu'est-ce que je connais dans cet arbre qu'est-ce que je connais dans cet arbre quand j'ai pas fait de retournements qu'est-ce qu'on considère cette composition*

62 *Qu'est-ce que tu mettrais ici comme probabilité*

Par rapport à

A P de A qu'est-ce que tu mettrais là au début

Il y a autant

Autant çà veut dire un cinquième un cinquième un cinquième un cinquième et un cinquième alors le celui-là là où est-ce que tu mettrais cette probabilité dans ton arbre celui que j'ai entouré tu le mettrais ici et puis si j'ai cette composition la probabilité d'obtenir N c'est exactement un et puis je peux continuer ici je mets un quart ici je mets un demi ici trois quarts est-ce que vous pouvez maintenant déterminer la probabilité que si je retourne la bouteille j'obtiens orange

Un demi

Bien sûr

Çà serait zéro plus un quart plus trois quarts plus un

Tu dis additionner tout çà et çà fait combien zéro plus un demi plus trois quarts plus un est-ce que c'et quelque chose qui vaut plus qu'un

Mais il faut le diviser par

Par divisé par cinq

C'est comme çà que tu vas raisonner alors je vais essayer de te mettre un peu plus sur la voie ici quelle est la probabilité

Zéro cinq c'est zéro cinq

Quelle est la probabilité d'obtenir quelle est la probabilité d'avoir la composition la composition et d'obtenir orange si je retourne la bouteille

C'est zéro

C'est zéro quelle est la probabilité d'obtenir A pardon que la composition soit orange B pardon et que la boule soit orange çà va être quoi la probabilité çà va être quoi wé comment tu fais

Un quart fois un cinquième

Un quart fois un cinquième alors je vous rappelle je vous rappelle à tous la question c'est la suivante comment vous faites pour calculer et la vous me dites que c'est P de B multiplié par P de O sachant B ce P de O sachant B c'est un quart P de B c'est un cinquième et c'est comme çà qu'on fait donc ceci est égal à un quart fois un cinquième quelle est la probabilité de P de C alter O

Un dixième

Un cinquième fois un demi quelle est la probabilité ici un cinquième fois trois quarts ici un cinquième et maintenant grâce à çà quelle est la probabilité d'obtenir O

Zéro cinq

La probabilité d'avoir O c'est quoi c'est çà plus çà plus çà plus çà plus çà c'est-à-dire zéro plus un vingtième plus un dixième plus trois vingtièmes plus un cinquième et tout çà çà fait quoi çà va faire regardez un vingtième plus trois vingtièmes çà fait quatre vingtièmes c'est la même chose que un cinquième et un cinquième plus un cinquième çà fait deux *cinquièmes donc deux cinquièmes qu'est-ce que j'ai oublié d'additionner un dixième çà va faire combien çà va faire un demi alors autrement dit on a fait beaucoup de calculs on a fait beaucoup de calculs et on arrive à quelque chose qu'est assez intéressante c'est que la probabilité d'obtenir l'événement O c'est-à-dire tant que on ne sait rien tant qu'on ne sait rien si je ne sait rien quelle est la probabilité d'obtenir orange c'est à priori un demi quelle est la probabilité d'obtenir noire un demi bon maintenant je sais donc la seule chose j'ai pas encore fait mon retournement avant on peut dire que la probabilité d'obtenir orange si je n'ai aucune information la probabilité d'obtenir orange et la probabilité d'obtenir noire c'est un demi ce c'est qu'on vient de calculer bon alors je sais çà maintenant que je fais ce retournement maintenant que je fais ce retournement que j'ai obtenu orange maintenant que j'ai fait ce retournement j'ai obtenu orange qu'est-ce que je cherche je cherche les différentes probabilités est-ce que vous pouvez lui donner un nom à cette probabilité le premier point d'interrogation là bas je cherche quelle probabilité donc la probabilité d'obtenir l'événement A sachant que c'est la boule orange qui est sortie est-ce que çà cette probabilité je la connais bah oui elle vaut zéro donc là je n'ai pas de problème celle-là vaut zéro la suivante le deuxième point d'interrogation là çà c'est quoi Sylvain ce que cherche c'est quoi regarde on sait que la boule qu'on vient de tirer elle est orange et on cherche la probabilité que la composition çà soit la composition B comment çà se écrit en langage de probabilité P de B sachant O alors la probabilité qu'on cherche c'est la probabilité de B sachant O qu'est-ce que je connais moi je connais le zéro virgule deux c'est quoi ce zéro virgule deux c'est quoi ce zéro virgule deux c'était le probabilité comment on l'appelle ce zéro virgule deux avant de retourner la bouteille la probabilité d'obtenir le d'avoir comme composition la composition B donc P de B il vaut zéro virgule deux un cinquième qu'est-ce que je connais d'autre qu'est-ce que je connais d'autre je connais cette je connais une chose ce que je vous disais tout à l'heure je connais la composition de la bouteille B bah je connais quand la composition est B je connais la probabilité d'obtenir quoi d'obtenir orange çà c'est ce que je connais çà vaut un quart et çà s'écrit comment la probabilité est égal à un quart ce quoi ce un quart qui apparaît là dedans c'est la probabilité d'O sachant B alors regardez notre problème on se dit on cherche la probabilité de B sachant O sachant qu'on part d'un principe avant d'avoir une information qui nous est transmise quand on retourne la bouteille j'ai a priori je lui attribut une certaine probabilité à l'événement à la composition B je lui ai affecté un certain nombre de pièces de monnaie et puis je connais quoi je connais O sachant B je cherche çà connaissant tout çà alors on va voir que maintenant bien avec tout ce qu'on sait faire tout ce qu'on a apprit à faire en probabilité avec les conditionnelles çà c'est facile à trouver c'est quoi P de B sachant O allez on applique les formules qu'on connaît*

P de B inter O

72 *Oui P de B iter O sur P de O ah mais il y a une truc que j'ai pas mis là dedans mais qu'est-ce qu'on a caclulé avant d'avoir retourné la bouteille qu'est-ce qu'on connaissait aussi on connaissait une probabilité d'obtenir orange et quelle est la probabilité d'obtenir orange c'était un demi avec le peu d'information qu'on avait on avait çà on avait et on avait çà bon voilà je sais que P de B sachant O est égal à çà est-ce que maintenant connaissant tout çà je peux pas le trouver le P de B sachant O alors une idée*

Bah si P de B sachant O c'est Zéro deux fois

On essaie de raisonner simplement par des lettres et après on remplacera

P de B fois P de O sachant B

74 *Wé alors tu me dis que çà c'est égal à P de B fois P de O sachant B et le tout divisé par P de O pourquoi et bien parce que je sais que P de B …O tu sais que il y a deux trucs que tu sais tu sais cette formule et tu sais aussi que P de B …O c'est égal à P de B multiplié par P d'O sachant B autrement dit je vais dire en mettant une signification à O et B quand je* *pars d'un principe j'ai pas encore retourné j'ai pas encore retourné la bouteille j'ai pas encore retourné la bouteille quand j'ai pas encore retourné je peux dire que la probabilité que la composition çà soit B c'est de zéro deux la probabilité que d'obtenir une pièce orange si je la retourne sachant que c'est la composition est B elle vaut un quart la probabilité d'obtenir orange elle vaut un demi de tout çà qu'est-ce que je peux en déduire je peux en déduire la probabilité que çà soit la composition B et que j'obtiens orange*

75 *Je peux obtenir la probabilité de P de B inter O cette probabilité P de B inter O elle je remplace dans cette formule et je trouve que la probabilité de B sachant O c'est quoi on remplace maintenant Jenny tu me dites*

Wé zéro deux fois un quart

Fois un quart

Divisé par un démi

Par un demi et çà fait combien

Çà fait zéro virgule zéro cinq divisé par zéro cinq je crois

Alors çà fait zéro virgule zéro cinq

divisé par zéro cinq

et diviser par un demi çà fait la même chose que multiplier par deux et çà fait zéro virgule

Un

Donc qu'est-ce que tu viens de trouver qu'est-ce qu'on vient de trouver on vient de trouver cette probabilité elle est de zéro virgule un je prends le tableau

76 *On est ici la probabilité d'obtenir A c'est zéro celle-là c'est zéro virgule un bon comment je pourrais faire pour calculer la probabilité cette probabilité que je cherche ici comment je l'appelle c'est la probabilité de quoi la probabilité de C sachant O je fais les mêmes calculs je fais les mêmes calculs et on obtient la probabilité alors on va ne pas le faire c'est là qu'on va prendre l'ordinateur il va arriver pour nous aider parce qu'il va faire tous ces calculs à la fois néanmoins j'essaie de généraliser le calcul qu'on a fait sur un que m'a permis de trouver la probabilité de B sachant O en fonction de ce qu'on connaissez alors je veux dire au lieu de si j'applique ce calcul à n'importe quelle de ces cinq compositions possibles*

77 *La probabilité d'avoir une des compositions sachant que la pièce qu'on tiré bah la boule qu'on obtient lors d'un retournement est orange si j'applique la formule çà donnerai çà P de O sachant K sur P de O fois P de K pour trouver c'est çà ce que j'ai fait avec B je le fais avec A je le fais avec C D et E çà ne change rien alors dans cette formule je vous demande d'observer une chose c'est que cette formule elle est de géométrie çà c'est quoi le P de K le P de B qu'on avait fixé à zéro virgule deux c'est notre état de croyances c'est notre état de croyances avant d'avoir retourné la bouteille et bon on se dit voilà la probabilité que ce soit la composition B c'est zéro virgule deux mais avant d'avoir retourné la bouteille*

78 *Et puis qu'est-ce que c'est qu'est-ce que c'est ce quotient là et bah c'est quotient il va changer c'est un quotient qui va déterminer le changement de probabilité parce que avant que je retourne la bouteille j'ai affecté un certain nombre à B j'a affecté une certaine probabilité à B et puis la probabilité que je cherche B sachant O c'est quoi cette probabilité et bien c'est la probabilité c'est après l'arrivée de l'information la nouvelle croyance que j'ai saisissez çà correspond à vos pièces vous avez vos pièces une fois que j'ai retourné la bouteille vous avez changé vous avez évolué la distribution de vos pièces vous aviez un état de croyance initiale que vous mesurez avec cette probabilité vous le faites changer et vous obtenez un état de croyance final après l'arrivée de la nouvelle information ok*

79 *Et bien çà c'est qu'on appelle c'est la nouvelle probabilité c'est la probabilité de K sachant O donc on va arrêter là chacune de ces probabilités j'ai simplement commenté le passage de la première à ma deuxième ligne chacune de ces probabilités qu'on cherche j'ai un état*

de croyances j'ai un état de croyances dont je n'ai pas retourné ma bouteille il est faible mais la croyance on dit on ne sait rien donc on l'affecte on l'affecte la même probabilité et cet état de croyance elle va je le fais évoluer vers une nouvelle probabilité je connais la probabilité de B et après je cherche la probabilité de B sachant O et puis si en suite une boule noire et bien

80 *Maintenant c'est quoi mon état de croyances c'est le zéro virgule un je crois en cette composition avec une probabilité de zéro virgule un et puis si j'obtiens noire je vais chercher une nouvelle probabilité et elle s'écrirait comment et bien elle est la probabilité d'obtenir B sachant que mon tirage est noire et cette probabilité j'utiliserais exactement la même formule pour la calculer vous avez vu bon alors c'est la pause cinq minutes et après vous travaillerez avec l'ordinateur qui nous aidera avec ces calculs*

81 *Vous avez sur le bureau un fichier que s'appelle bouteille vous démarrez*

82 *Bon alors le fichier Excel il va prendre les calculs qu'on a commencé qui va appliquer les formules qu'on a déjà qu'on a vu ensemble çà marche pas*

83 *Dans la feuille que je vous donne vous avez le mode d'emploi de ce fichier je vous explique comment pouvoir faire et en suite une fois que vous avez compris comment marche le fichier vous allez répondre au questionnaire que je vais vous donner alors on va commencer quand même ensemble pour que vous compreniez bien donc on est pour le moment on n'a fait aucun tirage on n'a fait aucun tirage on a dit qu'a priori quelle probabilité on a affecté à chacune des compositions c'était la probabilité zéro virgule deux donc ici il suffit avec les boutons de mettre*

84

85 *Donc ce que vous voyez apparaître lorsque vous avez mis des probabilités à vingt pour cent le total est de cent pour cent ici et ce que vous voyez apparaître un graphique qui vous donne et bien la probabilité que vous affectez à chacun à chacune des compositions alors nous qu'est-ce qui s'est passé et bien ce qui s'est passé qu'on a obtenu en premier on a obtenu orange c'est bien çà donc je vais rentrer O ici je rentre O et maintenant avec le principe qu'on a vu tout à l'heure*

86 *Vous voyez que là j'ai choisi orange la probabilité la probabilité de B la probabilité de B sachant qu'on a obtenu orange la probabilité de B est devenu zéro virgule un qu'on a calculé tout à l'heure de la même manière suivant le même principe les autres probabilités elles ont été calculées ici bon alors qu'est-ce qu'on a obtenu en deuxième on a obtenu noire alors ce que je vous demande d'observer*

87 *A chaque fois on va voir l'évolution des probabilités qu'on voit apparaître sous la forme de diagramme à trois dimensions donc quand on est ici on retrouve c'est à partir de maintenant que votre travail il va il va commencer c'est-à-dire que vous allez déjà interpréter vous allez comparer ces résultats des résultats qu'on a justifié grâce à ce qu'on sait donc vous allez comparer vos croyances initiales que vous avez répondu et après des que vous auriez fait cette partie moi je tirerai d'autres je ferai d'autres tirages je ferai d'autres retournements donc déjà maintenant là où vous êtes vous pouvez commencer le questionnaire çà commence par les pièces de monnaie et vous pouvez commencer à réfléchir à répondre à comparer (1h21m)*

Cette feuille est pour les probabilités les probabilités qu'on a calculé tout à l'heure donc les probabilités qu'on a calculé tout à l'heure là vous les

	mettez ici vous les notez là			
99		*Donc la feuille numéro deux c'est çà vous justifiez çà c'est pour mettre les probabilités et après là l'ordinateur applique la formule les calculs qu'on a fait et vous allez commenter les calculs*		
	Il y a en deux lignes en dessous c'était plus ou moins aléatoire		*Là c'est le récapitulatif des probabilités et là ce sont les calculs qu'on a fait et vous avez les résultats qu'a fait l'ordinateur*	
	La probabilité dans cette feuille là			
	Ah oui oui			
	Là vous mettez les probabilités celles qui apparaissent ici c'est qu'il y a c'est d'essayer de le justifier pourquoi pour celle-là par exemple là le P de A sachant O là			
			Dans la feuille numéro deux elle corresponde à çà et il suffirait mettre les couleurs l'ordinateur fait les calculs	
		Je ne sais pas comment le justifier		*Donc vous avez répondu vous regardez vous remplissez çà*
	Çà ce sont les probabilités en suite premier tirage la probabilité pour celle-là est zéro et après celle-ci va être			**Moi j'essaie de voire comment fonctionne le truc**

zéro aussi voilà			
P de A sachant O c'est	*Alors tu t'en souviens comment on a fait celui-là c'est quoi celui-là alors moi je t'ai donné des informations quand on est là on a dit que l'a probabilité de O c'est un demi et que la probabilité de N est un demi et cette probabilité qu'on cherche c'est la probabilité d'obtenir B sachant qu'en retournant la bouteille on a obtenu orange là tu as la formule*		*Bah si tu veux au-delà bah ne fait il faut que tu utilisez la chose suivante après un tirage çà c'est a priori ce sont tes probabilités avant après premier tirage tu as le l'A qui vaut zéro le B qui vaut deuxième tirage et puis en suite troisième tirage tu vois l'évolution des probabilités là tu as zéro parce qu'on n'a fait d'autres tirages c'est pour çà*
C'est zéro	**Et P de B inter O**		**A d'accord**
	La probabilité de B sachant O la probabilité de O sachant B regarde là la formule		*Pour le moment on n'a que trois*
Divisé par P de O je crois si si par P de O ou par P de A c'est çà P de A inter O çà va être P de A fois P de O sachant A sur P de O oui c'est de P d'inter sur P de O et çà P de O inter A fois P de O sachant A		*C'est là justifiez quelques calculs et comparez avec ce que vous avez fait c'est intéressant comment vous justifiez regardez la composition la plus éloignée de la distribution et essayez de voir pourquoi est-ce qu'elle est très éloignée et comparer avec votre raisonnement de départ*	*Les deux graphiques te donnent la même*
Parce que P de O sachant A c'est zéro			**Comment**
Mais il faut justifier mais tu as pas besoin de justifier que P de O sachant		*Et là la probabilité qu'on cherche est la probabilité de B sachant qu'on a*	*Çà et çà te donnent la même information et là vous justifiez quelques calculs*

A c'est zéro çà on le connaît

trouvé qu'on a obtenu qu'est-ce qu'on a obtenu on a obtenu une boule orange donc c'est la probabilité de B sachant O qu'on l'obtient avec la probabilité de O sachant B et à chaque fois la formule

vous comparez avec ce que vous aviez fait c'est intéressant essayez d'expliquer indiquez en quelle ligne votre distribution est plus éloignée et puis et essayez pourquoi pourquoi elle est très éloigné et qu'est-ce que vous aviez fait comme raisonnement

83 *Je vous ai mis pour ceux qui ne se souviennent pas la formule pour que vous essayez de l'appliquer dans vos calculs çà vous permettra de bien justifier le calcul qui permet de faire évoluer de faire évoluer les probabilités de faire évoluer nos croyances*

84 *Bon alors on va continuer je vais faire d'autres retournements donc on avait dit qu'on avait obtenu donc c'est bien çà premier retournement j'ai obtenu O orange deuxième retournement j'obtiens noire et je vais continuer et comme çà vous pourrez une fois que vous auriez fait les calculs on va voir*

85 *Troisième retournement j'obtiens noire quatrième retournement*

Orange

J'obtiens orange cinquième

Noire

Noire sixième

Orange

Orange c'était le sixième septième orange

86 *Huitième*

Orange

Noire

Neuvième orange et orange voilà bon alors n'oubliez pas de justifier vos calculs et puis vous répondez aux questions

87 *Alors comment est-ce que tu l'interprètes par rapport au problème qui était qui était c'était relatif à la composition dans notre problème c'était d'estimer d'estimer la composition çà tu l'interprètes comme la probabilité que la composition C sachant qu'on a obtenu cette boule c'est ta probabilité c'est ton état de*

Il doit y avoir une noire et trois oranges tu va voir il y a eu beaucoup d'oranges

	croyance en C est celui-là et par rapport à la bouteille tu peux dire tu vois çà et tu peux dire que la composition çà risque d'être celle-ci avec une probabilité de çà même si tu ne sais rien ce qu'il y a dedans tu le sauras jamais mais tu arriver à dire quand même avec une certitude de çà que la composition est la C d'accord et ici tu justifiez les calculs			
	Bon alors moi je suis plutôt pour la C	**On va voir si çà fonctionne**		*Alors*
	On met qu'on a moins de chances d'avoir une orange qu'une noire	**Tu es en forme aujourd'hui**		**C'est quoi ces chiffres**
	Nous on avait mis quatre virgule quatre pour cent à la première			*Tu arrives pas à les interpréter on essaie de les interpréter ensemble Jeremy*
89			*Là tu as calculé la probabilité de la composition C sachant qu'on a eu une boule orange*	**Moi je l'ai fais**
			Là c'est P de O sachant B c'est la probabilité que tirer une boule orange sachant que la composition est la B	*Alors toi tu les as interprété travaillez à deux par exemple notre problème c'est quoi c'est d'estimer la composition de la bouteille là tu as fait dix fois le mâchant qu'est-ce que tu peux dire*
	Parce qu'en fait on avait mis qu'est-ce qu'on a avait mis on avait mis cinquante pour		*La tu arrives à dire que la probabilité de la compositions C est de cinquante cinq pour cent et regarde*	**Bah qu'il y a la la que c'est C**

cent de probabilité pour le billes orange et cinquante pour les noires		*là là on hésite entre les deux on hésite entre les deux et est-ce qu'on peut dire autre chose alors là c'est autre chose on estime*	
Çà c'est bon	*Alors vous comprenez bien le truc finalement*		*Mais quelle est la probabilité*
Bon mais là il faut justifier aussi	**Pour moi c'est celle-ci**		
Oui mais des trucs de hasard justifier çà c'est un peu bizarre	*Ok tu dis NNOO tu dis que c'est la composition C et après avoir fait dix retournements tu arrives c'est quelle probabilité pour cette composition que çà soit cette composition*		*Tu arrives à dire que la probabilité que ce soit C c'est de cinquante trois virgule comment tu l'interprètes est-ce que tu sais si c'est la C*
Dans quelle ligne	**Je crois que c'est celle-là**		**Bah non**
	Wé		*Voilà mais tu arrives alors on oublie tous ces calculs quand tu vois çà tu fais dix tirages comme çà qu'est-ce que tu peux dire*
Pourquoi tu mets pas oui et juste deux	Parce qu'en fait il y a plus d'oranges la D je dis il y a plus d'oranges que de noires		**Qu'il y a plus d'oranges**
	Oui d'accord mais on va essayer de une autre manière regarde c'est quoi la proportion d'oranges par rapport aux noires deux quatre six si normalement c'était cette cette composition-là tu aurais ...pour cent pour celle-là c'est quoi ton calcul là tu		*Plus d'oranges mais tu dirais quel type de composition*

	vois quelque part avec tout ce qui se passé avant parce que cette information que tu as eu avant elle te donne tu vas avoir tu vas privilégier celle-ci par rapport aux autres		
De la colonne deux	**Je sais pas comment l'expliquer après dix retournements on va dire que**		**La D**
Attends attends çà c'est quoi çà c'est les lignes çà c'est les colonnes et la ligne trois c'est çà	*Mais qu'est-ce qui te dis que c'est çà la composition*		*La D mais comme il apparaît beaucoup plus tu vois la proportion ici elle ne corresponde pas on pourrait dire c'est entre les deux quoi on hésite entre les deux mais on hésite mais comme çà tu peux pas faire autre chose qu'avoir une idée alors que là on fait autre chose on estime caremment peut être est celle-là je n'en sait rien mais en fonction des billes tirées on dit qu'est celle-là avec une probabilité de que tu as là*
Et pour les colonnes	**Bah les probabilités quoi**		**Mais ce sont des chiffres complètement aléatoires**
Moi j'ai pas pris les colonnes	*Les probabilités tu sais que c'est cette composition avec telle ou telle probabilité*		*Ces chiffres ils traduisent ils vont mesurer le hasard le hasard il va être toujours là mais avec çà tu peux quantifier le hasard je peux affirmer que en ayant ces tirages*

			j'ai cinquante cinq virgule huit pour cent de chances que se soit cette composition
Ligne deux je ne comprends pas dans quelle ligne tu es	**A cinquante cinq pour cent**		*Mais si on fait plus de tirages*
Çà	*A cinquante cinq pour cent si tu veux tu le donnes en termes de croyances tu as une croyance à cinquante cinq pour cent çà vaut dire qu'à cinquante cinq pour cent de chances que soit celle-ci*		*Plus on fait tu vois regarde regarde avant tu vois qu'avant on avait même des probabilités qui étaient tu vois c'était C à soixante un pour cent*
	Il y a un risque		**Mais celui-là il monte**
Orange	*Tu vas avoir un risque*		*Après on a eu d'autres informations çà a un petit peu relativisé cette information quoi tu comprends*
Qu'est-ce que c'est une composition ah voilà la composition c'est C quoi	**Le risque il es partie en deux probabilités quoi B et D**		**Par contre c'est que aussi ah non non c'est normal c'est normal**
De tout çà je ne comprends rien	*Ah oui bien sur tu as pour celle-ci cinquante cinq pour cent et pour les autres le reste*		*Tu vois le fait qu'une orange qu'arrive ne plus ici c'est très très difficile à avoir quand même mais çà va quand même plutôt ce qu'on avait avant si tu as une orange qu'arrive encore çà va te faire baisser celle-là parce que là tu avais quand même très forte croyance celle-là si tu as une orange qu'arrive en plus çà va te faire baisser et çà va te*

		faire çà va te faire augmenter celle-là el là elle augmente beaucoup
	Voilà çà	
Les billes indiquées	*Mais si tu dois prendre une décision c'est quoi qu'il y a dedans*	
Il faut mettre orange noire orange noire	**Bah oui**	
Attends si l'un des billes	*Tu sais que tu prennes un risque oui mais ce que tu as là c'est que tu arrives à mesurer ton degré de croyances alors je te dis regarde tous tes tirages là*	
Mais ce quoi çà change pas	**Moitié moitié quoi**	
Indiquez votre choix	*Là tu pourrais dire tiens c'est la C mais là tu arrives à mesurer avec précision*	
Attends je n'a pas de justification justifiez attends alors je ne sais pas quoi mettre	**Même quand on voit le graphique là**	
Si oui pourquoi	**Wé**	
çà m'énerve cette histoire de justifier	**On voit bien le C là qu'il tombe vraiment**	
Moi je justifies pas moi		
	…	
	Çà dépend du nombre de billes et tu le voit par rapport au graphique	
		Çà va vous avez réussi à bien remplir
		Çà me dis toujours rien les chiffres là

Qu'est-ce que vous faites là avec ta clé
Non mais en fait il avait laissé la protection
Bah allez on finit là alors vas-y pose moi tes questions
Que çà dépend totalement de ce que vous avez tiré pour moi çà pourrait bien être quelque chose d'autre c'est pour cela que je n'arrive pas à me prononcer là
Bah si tu veux imagine que çà soit réellement çà qu'est-ce qui apparaît le moins possible le moins resemblable c'est celle-là
oui
On est pas sûr que soit celle-là mais avec quelle probabilité est celle-là quatre pour cent pourquoi tu vois qu'avec cette information qu'on a on a une information que nous a été transmise
J'arrive pas à le mettre en mots pour le remarquer
Mais peu importe çà c'est pas très grave il suffit que tu arrives à bien comprendre tu as une information qui t'a été transmise et avec tout cette information qui t'a été transmise

t'arrives à dire que çà çà c'est sûr que çà peut avoir lieu on ne sait rien du tout moi je ne te dirai jamais ce qu'il y a dans la bouteille

Ah non

Ah bah non ah bah non parce que finalement ce que nous importe c'est pas de savoir ce qu'il y a c'est de trouver une méthode qui nous permet d'estimer avec quelle chance quelle probabilité donc tu peux affirmer tu comprend que cette idée d'avoir non pas tu auras jamais d'assurance tu peux pas affirmer ce qu'il y a dedans mais par contre je peux te dire qu'y a dedans il peut y avoir une de ces quatre possibilités mais pour chacune de ces quatre possibilités je peux d'afficher la chance qu'il y a pour chacun de ces quatre possibilités donc j'ai quelque chose un élément de réponse çà va pas me changer le hasard il y a toujours du hasard il y aura toujours du hasard tu aurais aucune certitude tu n'auras jamais certitude mais je peux préciser mon degré de certitude tu vois *l'idée et c'est çà l'idée de probabilité c'est une manière de comprendre cette probabilité la probabilité quand on fait un sondage on va jamais savoir qui va être élu mais çà peut te donner avec un çà peut te donner tu peux te dire qu'avec un certain risque et tu peux estimer*

Mais même les élections c'est pas aléatoire

Bah si c'est aléatoire quand tu prends milles personnes tu interroges les personnes

Les personnes sont

Si si tu veux le fait de prendre par exemple cinq mille personnes

Mettez vos noms aux feuilles

On a fait progresser la science

Même dans la science c'est pas du tout aléatoire rien n'est aléatoire dans tout çà il n'y a rien d'aléatoire rien d'aléatoire

Qu'est-ce que tu entendes par aléatoire

Et alors qu'est-ce qu'il y avait dedans monsieur

Donc tu penses que rien n'est aléatoire

Même dans la bouteille mais chaque boule

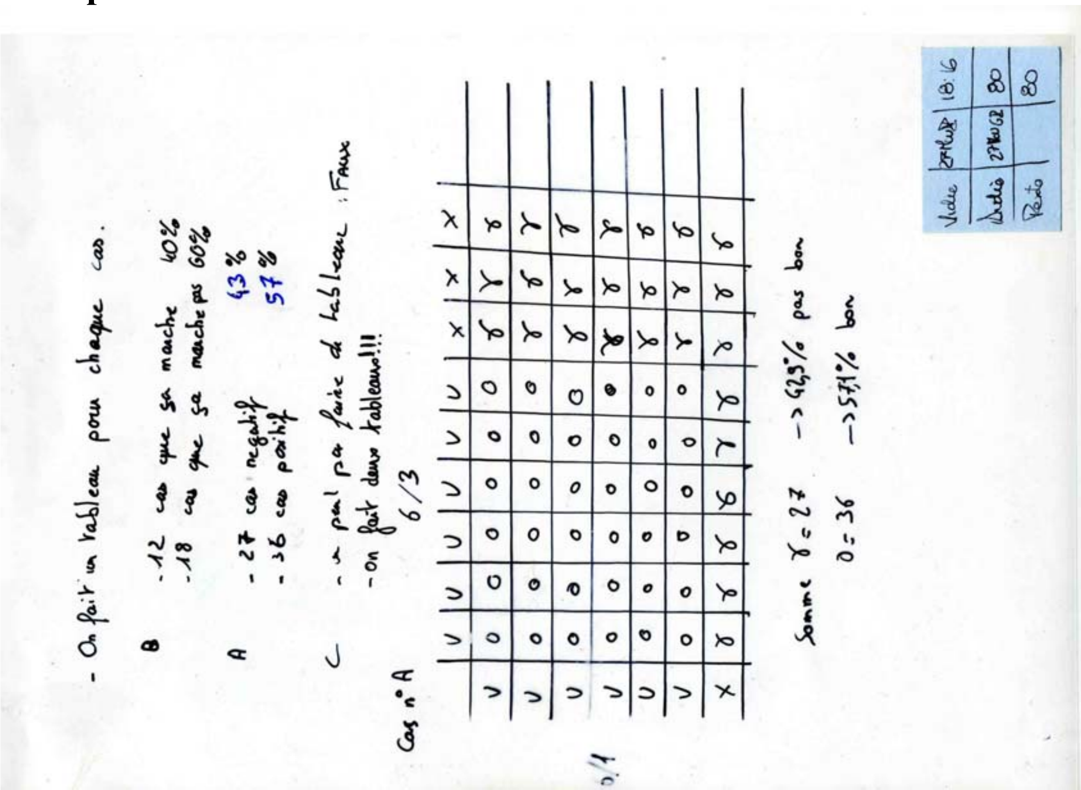
Tu as raison même si l'hasard n'existe pas *la probabilité*

Annexe III

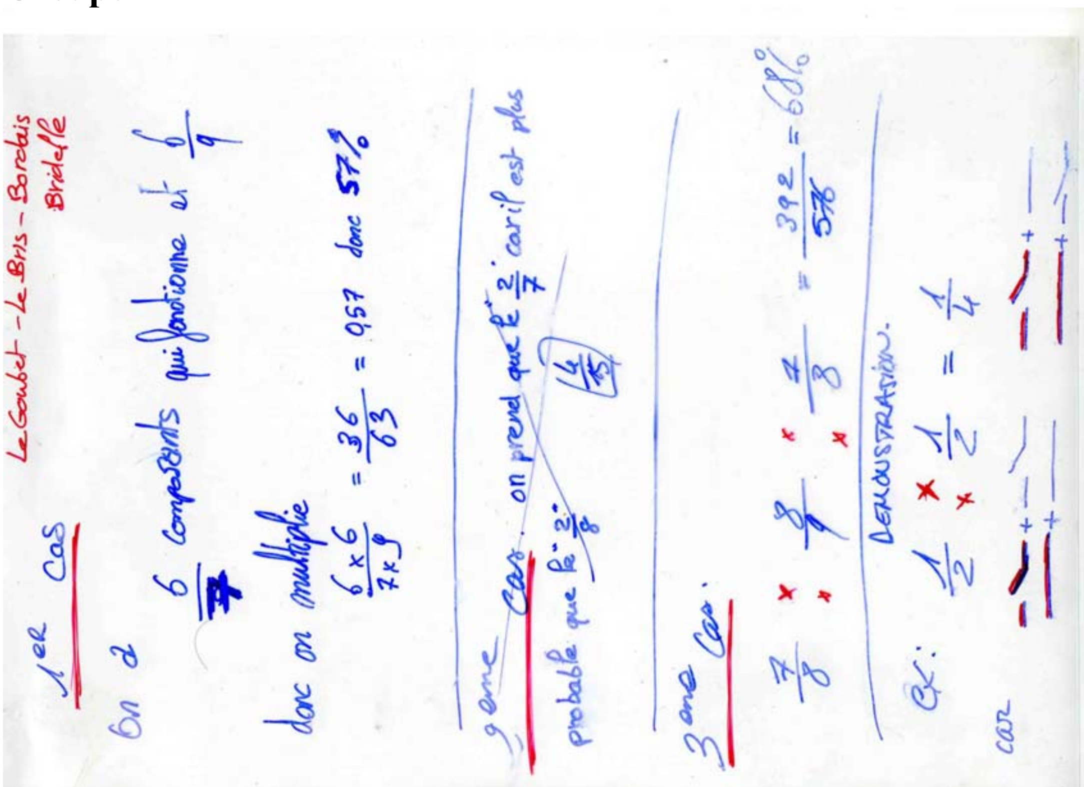
Autres documents

Transparents de présentations des élèves. Séance les circuits

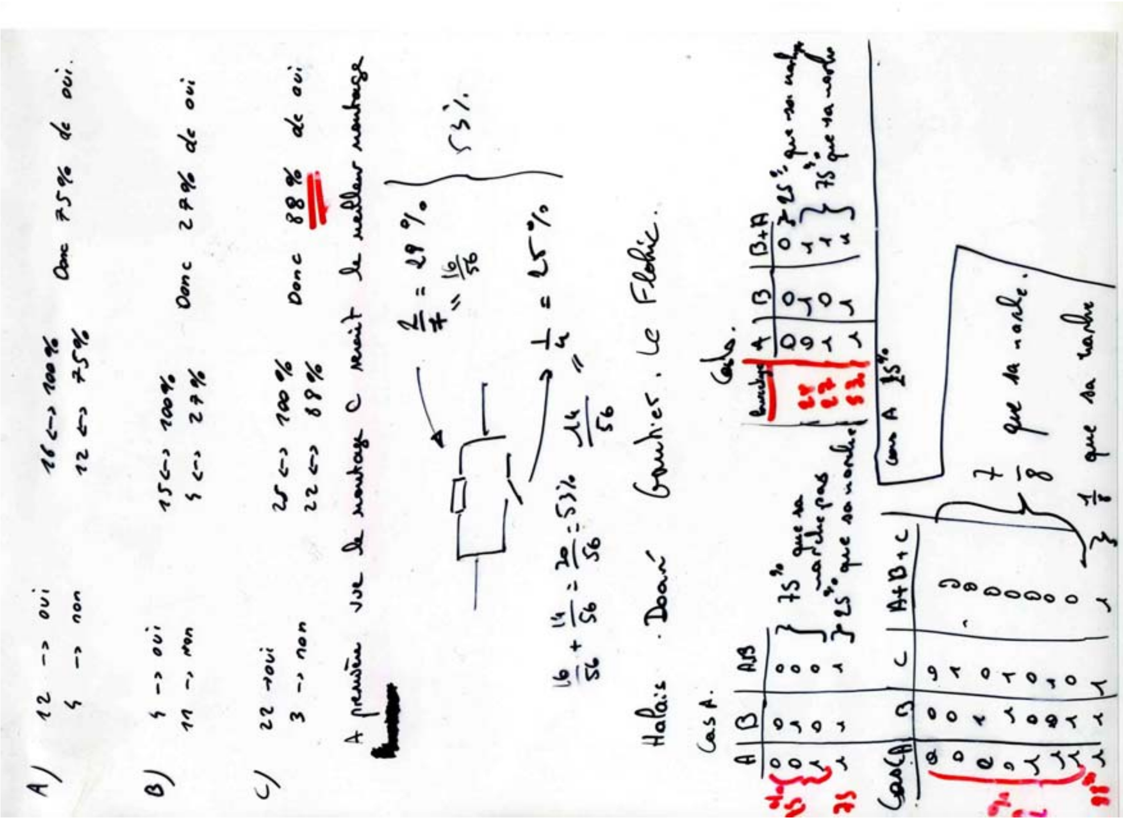
Groupe 1



Groupe 2



Groupe 3


A) 12 → oui ; 4 → non ; 16 ↔ 100 % ; 12 ↔ 75 % ; Donc 75 % de oui.

B) 4 → oui ; 11 → non ; 15 ↔ 100 % ; 4 ↔ 27 % ; Donc 27 % de oui

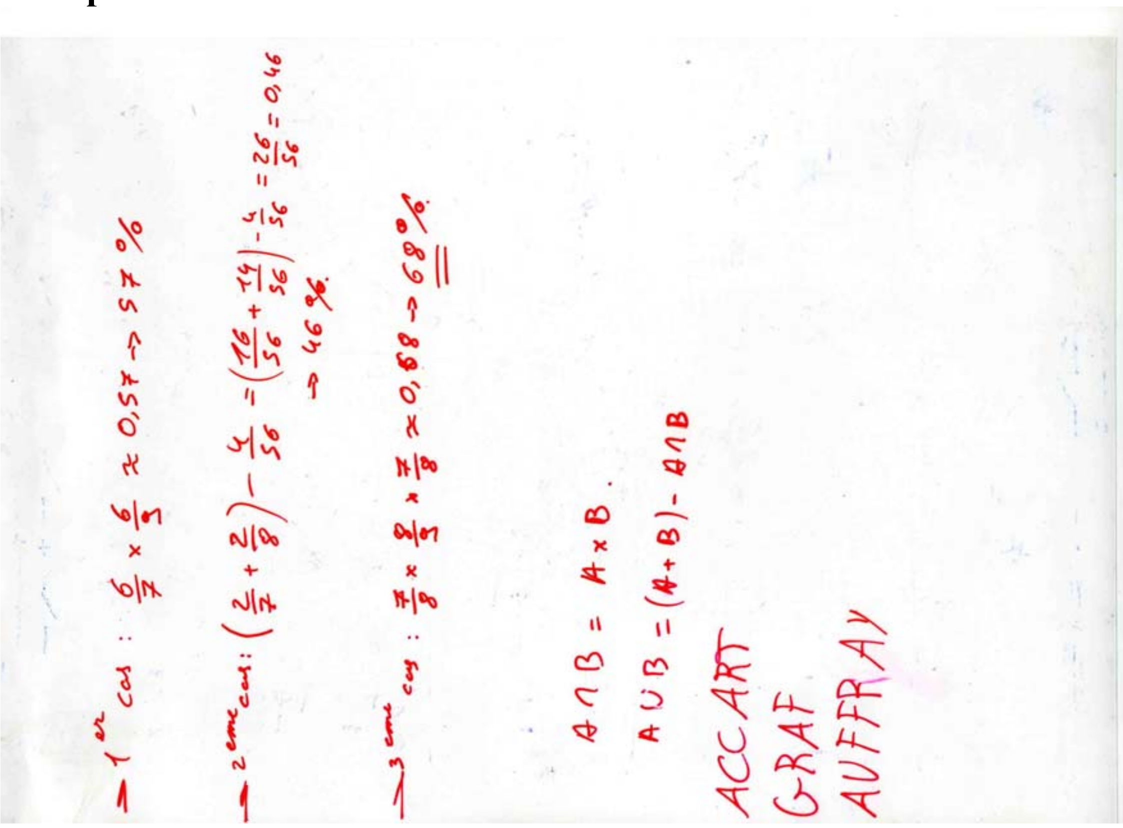
C) 22 → oui ; 3 → non ; 25 ↔ 100 % ; 22 ↔ 88 % ; Donc 88 % de oui

A première vue le montage C serait le meilleur montage

$\frac{2}{7} = 29\,\% = \frac{16}{56}$; $\frac{1}{4} = 25\,\% = \frac{14}{56}$; 53 %

Halais . Doan . Gauthier . Le Flohic.

Groupe 4


→ 1er cas : $\frac{6}{7} \times \frac{6}{8} \approx 0{,}57 \rightarrow 57\,\%$

→ 2ème cas : $\left(\frac{2}{7} + \frac{2}{8}\right) - \frac{4}{56} = \left(\frac{16}{56} + \frac{14}{56}\right) - \frac{4}{56} = \frac{26}{56} = 0{,}46 \rightarrow 46\,\%$

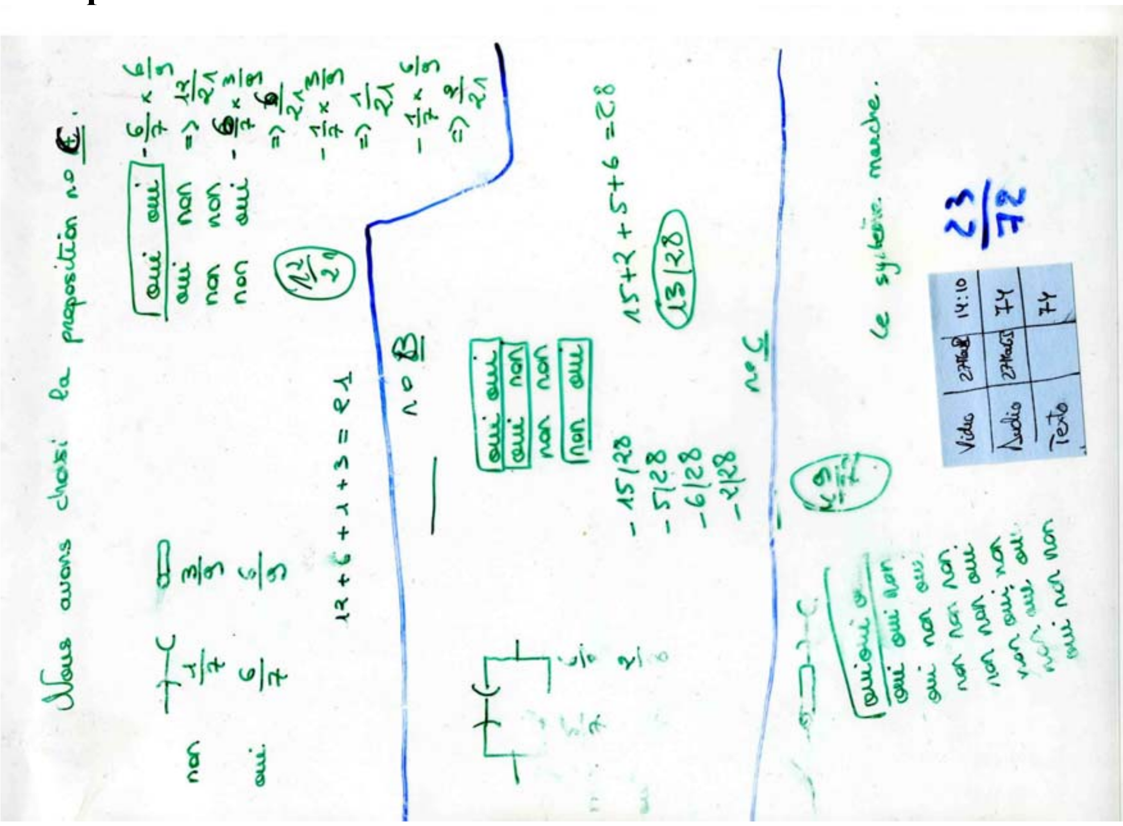
→ 3ème cas : $\frac{7}{8} \times \frac{8}{9} \times \frac{7}{8} \approx 0{,}68 \rightarrow 68\,\%$

$A \cap B = A \times B$.

$A \cup B = (A + B) - A \cap B$

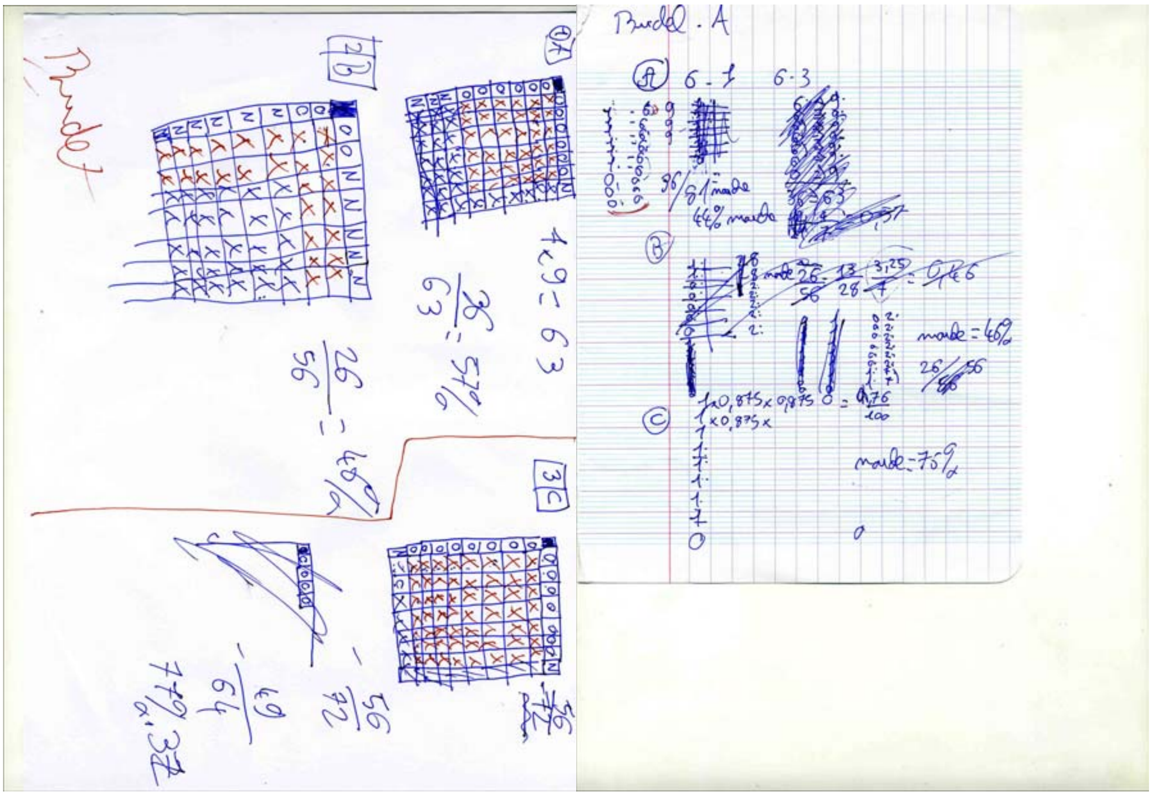
ACCART
GRAF
AUFFRAY

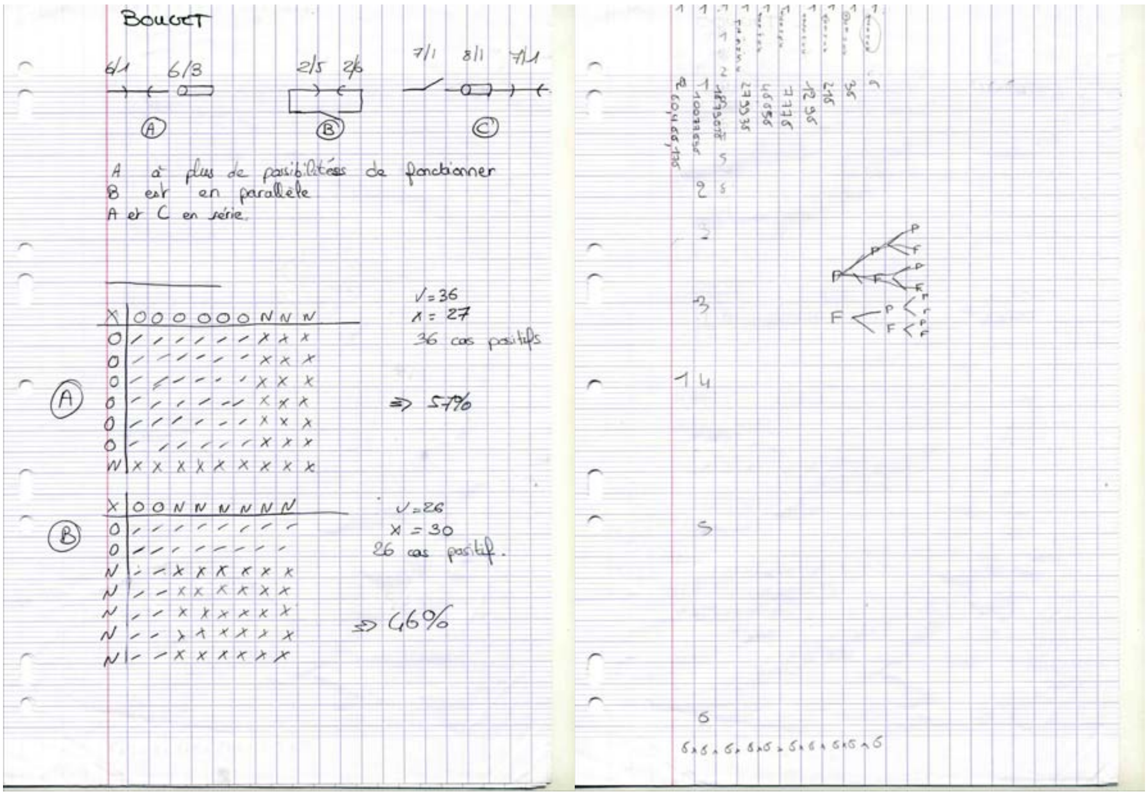
Groupe 5



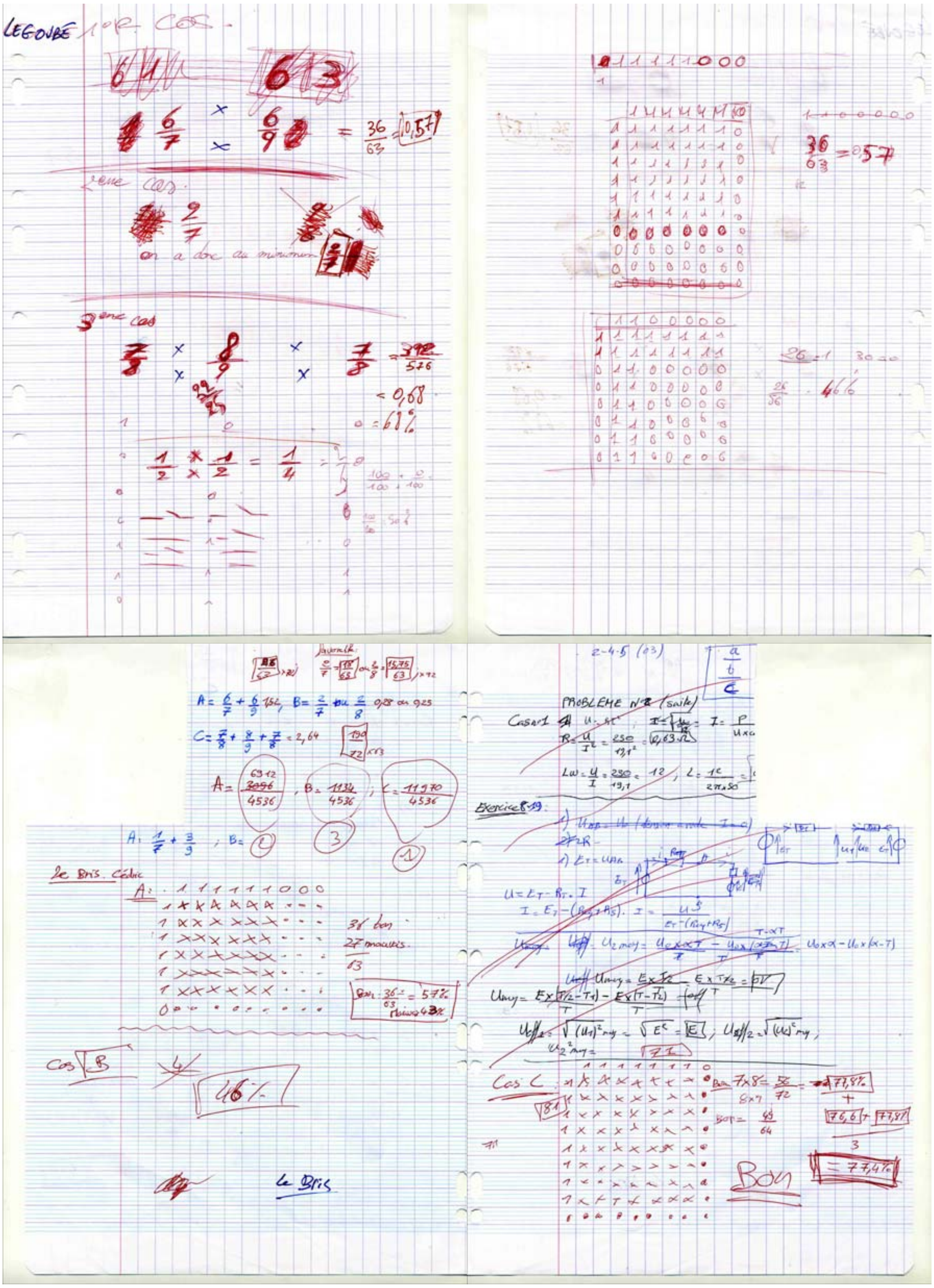
Feuilles de recherche. Séance les circuits

Groupe 1

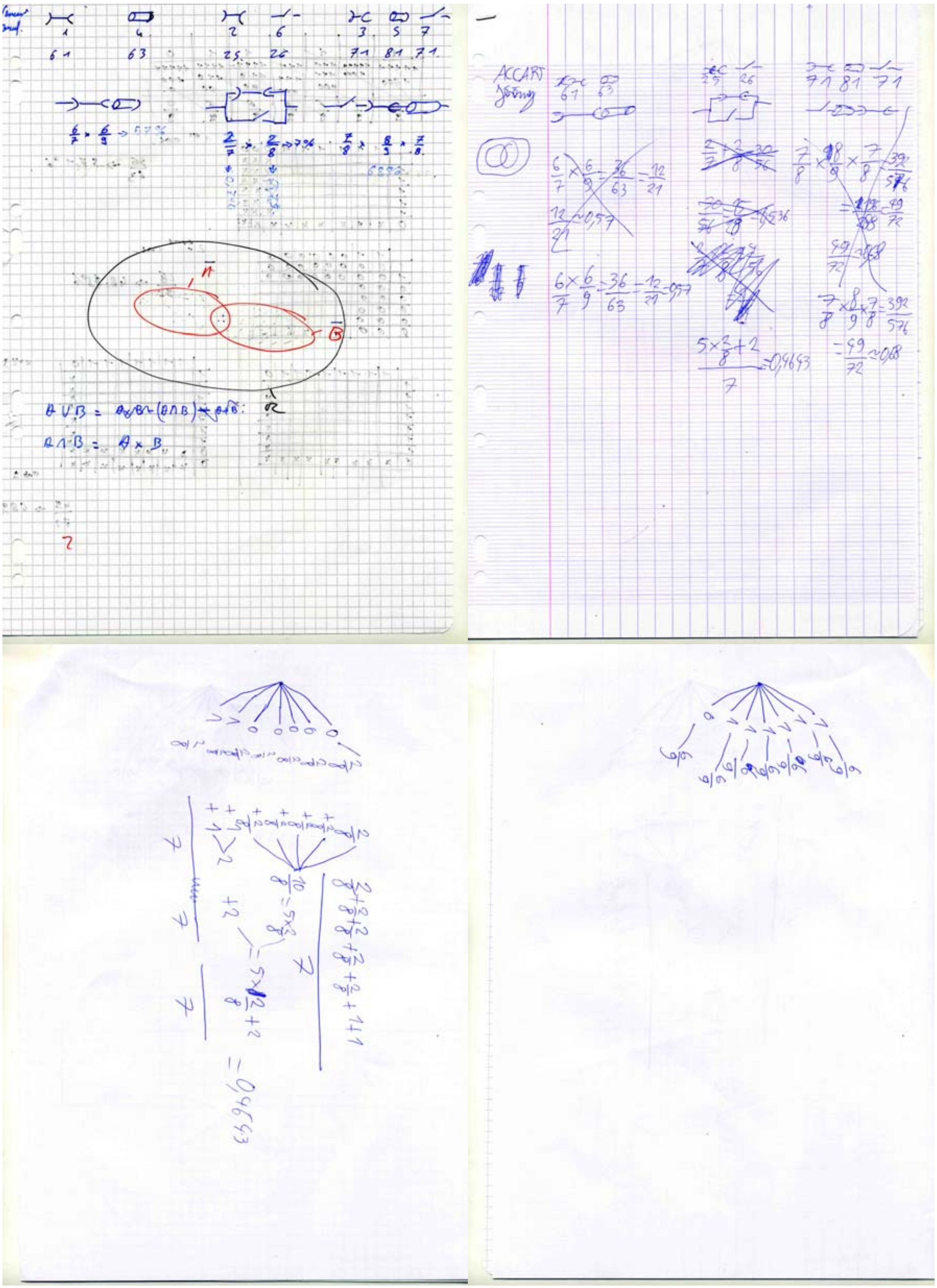


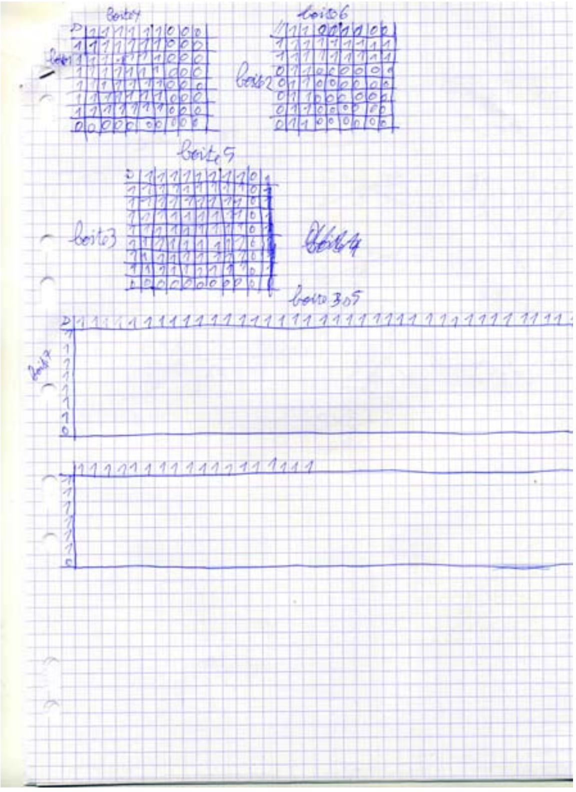


Groupe 2

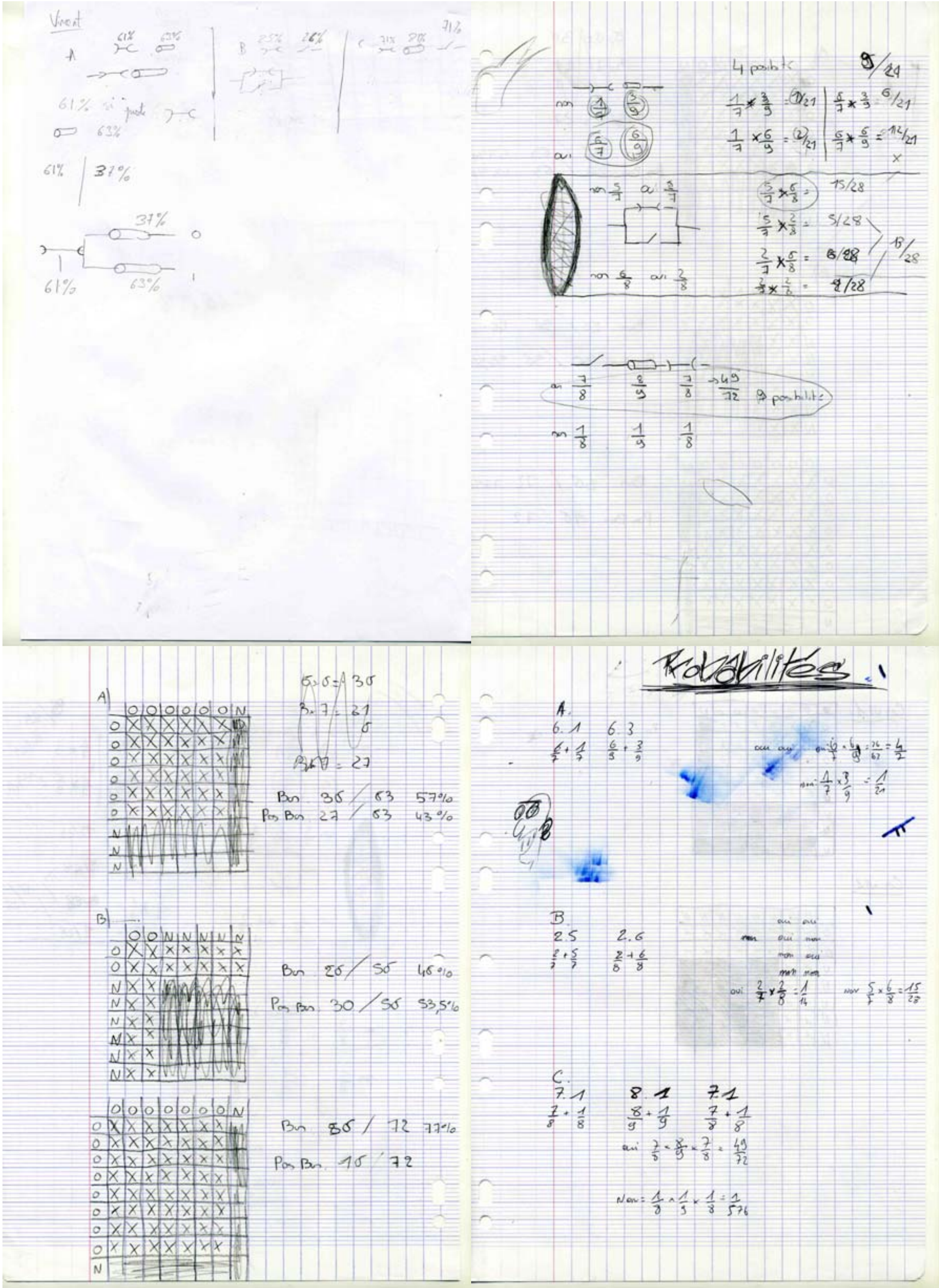


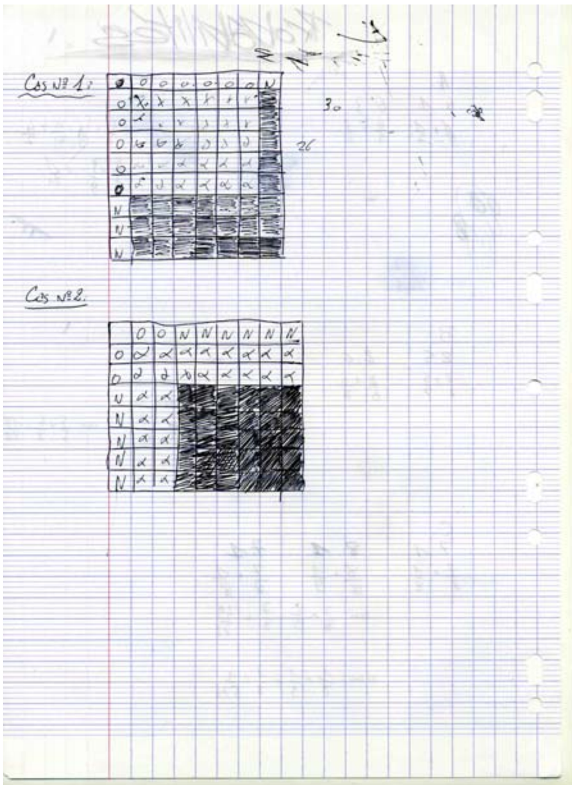
Groupe 4

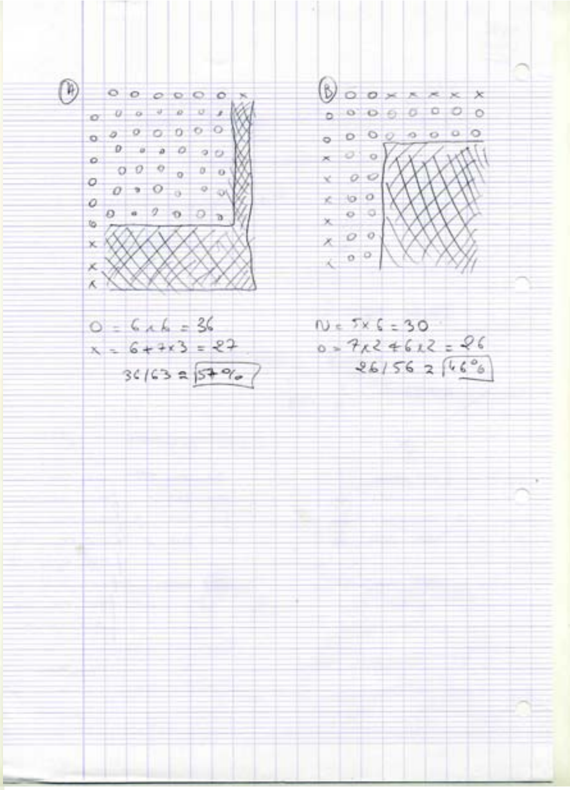


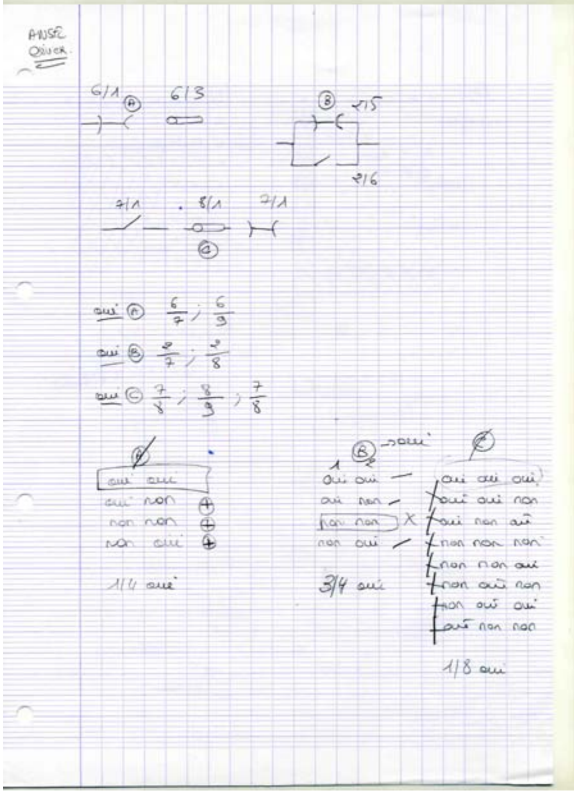


Groupe 5









Questionnaire. Séance la bouteille

Les pièces de monnaies

1) Recommencez le problème des pièces de monnaie du début, cette fois en termes de probabilités :

   a. Attribuez vos premières probabilités à l'aide des boutons jaunes (avant le premier retournement de la bouteille).
   b. Indiquez dans la première case verte claire la couleur de la bille du premier retournement (« n » pour noire, « o » pour orange).
   c. Indiquez dans la deuxième case grise la couleur de la bille du deuxième retournement.
   d. En quelle ligne votre estimation de la distribution de pièces est plus éloignée de la distribution de probabilités de la a feuille Excel « Bouteilles » ? A quoi attribuez vous cet écart ?

Les compositions

2) Demandez au professeur les retournements nécessaires pour compléter la colonne verte claire de la feuille Excel « Bouteilles » (dix retournements).

   a. Après ces dix retournements, estimez vous que la composition de la bouteille est de ? (cochez une seule option) :

      NNNN ☐
      NNNO ☐
      NNOO ☐
      NOOO ☐
      OOOO ☐

   b. Etes-vous persuadé de cette composition ?

   Oui ☐
   Non ☐

   c. Justifiez la réponse

…………………………………………………………………………………………………
…………………………………………………………………………………………………

3) Auriez-vous fait toujours le même choix que vous venez d'indiquer en 2)a), si l'on avait fait juste cinq retournements ?

   a. **Oui** ☐
      Non ☐

   b. Si oui, pourquoi auriez-vous choisi toujours la même composition ?

…………………………………………………………………………………………………
…………………………………………………………………………………………………

   c. Sinon, quel autre choix ?

      NNNN ☐

NNNO ☐
NNOO ☐
NOOO ☐
OOOO ☐

d. Comment expliquez-vous cette différence de choix ?
……………………………………………………………………………………………
……………………………………………………………………………………………

Les billes

4) Lors des dix retournements, le fait de connaître la couleur d'une bille a-t-il toujours modifié les probabilités précédentes au retournement de la bouteille ?

a. **Oui** ☐
Non ☐

Sinon :
b. Dans quelle ligne du tableau ? ………..
c. Quelle couleur de bille ? …………
d. Pour quelle composition ? ………..
e. Qu'est-ce que cela signifie ?
……………………………………………………………………………………………
……………………………………………………………………………………………

L'ordre

6) Après les dix retournements, rangez les compositions dans le tableau ci-dessous selon le critère suivant : plus vous êtes certain d'une composition, plus à gauche vous la placez.

a.

b. Justifiez votre rangement
……………………………………………………………………………………………
……………………………………………………………………………………………

7) Indiquez dans le tableau ci-contre les couleurs des billes obtenues lors des dix retournements (en respectant l'ordre d'apparition).

a.

Modifiez comme vous le souhaitez l'ordre et juste l'ordre d'apparition des billes. Indiquez ce nouvel ordre dans le tableau ci-dessous.

b.

c. Si l'ordre d'apparition des billes aurait été celui que vous venez d'indiqué en 7)b., changeriez-vous la réponse à la question 2)a) ? Vous pouvez utiliser l'ordinateur pour répondre cette question.

Oui ☐
Non ☐

d. Pourquoi ?

………………………………………………………………………………………………………
………………………………………………………………………………………………………

e. Si oui, indiquez votre choix dans le tableau ci-dessous :

NNNN ☐
NNNO ☐
NNOO ☐
NOOO ☐
OOOO ☐

f. A quoi attribuiez-vous ce nouveau choix ?

………………………………………………………………………………………………………
………………………………………………………………………………………………………